AUSTRALIAN HOUSING MARKET PROJECT

In this project will be divided into two main sections. Firstly, we will do research to find out the mean dwelling price, the number of residential dwelling in each state as well as the percentage change for housing price during Covid time until now. From these results, we can then analyze the housing market's healthy and trend in Australian's states.

The other section will compare the dwelling stock value owned by household and non household as well as how much changed regarding the stock value from 2020 quarter 1 to 2023 quarter 2.

First of all, let's import the data and do the insight check.

```
import numpy as np                  #For linear algebra
import pandas as pd                 #For working with dataset
import matplotlib
import matplotlib.pyplot as plt     #Visualisation
import seaborn as sns               #Visualisation
```

1. Data Exploration

Read the dataset and inspect its appearance.

```
df = pd.read_csv('/content/drive/MyDrive/ABS_RES_DWELL_ST_..Q.csv')
df.head()
```

	DATAFLOW	MEASURE: Measure	REGION: Region	FREQ: Frequency	TIME_PERIOD: Time Period
0	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-Q1
1	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-Q2
2	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-Q3
3	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-Q4
4	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2021-Q1

```
df.tail()
```

	DATAFLOW	MEASURE: Measure	REGION: Region	FREQ: Frequency	TIME_PERIOD: Time Period	OBS_V
625	ABS:RES_DWELL_ST(1.0.0)	5: Mean price of residential dwellings	3: Queensland	Q: Quarterly	2022-Q2	
626	ABS:RES_DWELL_ST(1.0.0)	5: Mean price of residential dwellings	3: Queensland	Q: Quarterly	2022-Q3	
627	ABS:RES_DWELL_ST(1.0.0)	5: Mean price of residential dwellings	3: Queensland	Q: Quarterly	2022-Q4	
628	ABS:RES_DWELL_ST(1.0.0)	5: Mean price of residential dwellings	3: Queensland	Q: Quarterly	2023-Q1	
629	ABS:RES_DWELL_ST(1.0.0)	5: Mean price of residential dwellings	3: Queensland	Q: Quarterly	2023-Q2	

The data looks fine except the OBS_STATUS: Observation Status and OBS_COMMENT: Observation Comment with most of column appears NaN (Not A Number). This is because not too much data in this report needs to be revised and there is no observation comment.

Now let's check the data types and shape.

```
df.dtypes
```

```
DATAFLOW                          object
MEASURE: Measure                  object
REGION: Region                    object
FREQ: Frequency                   object
TIME_PERIOD: Time Period          object
OBS_VALUE                        float64
UNIT_MEASURE: Unit of Measure     object
UNIT_MULT: Unit of Multiplier     object
```

```
OBS_STATUS: Observation Status       object
OBS_COMMENT: Observation Comment    float64
dtype: object
```

```
df.shape
```

```
(630, 10)
```

It shows that there are 630 columns and 10 rows in the data.

Let's go further to check if there are any missing data in other columns except OBS_STATUS: Observation Status and OBS_COMMENT: Observation Comment.

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 630 entries, 0 to 629
Data columns (total 10 columns):
 #   Column                           Non-Null Count  Dtype  
---  ------                           --------------  -----  
 0   DATAFLOW                         630 non-null    object 
 1   MEASURE: Measure                 630 non-null    object 
 2   REGION: Region                   630 non-null    object 
 3   FREQ: Frequency                  630 non-null    object 
 4   TIME_PERIOD: Time Period         630 non-null    object 
 5   OBS_VALUE                        630 non-null    float64
 6   UNIT_MEASURE: Unit of Measure    630 non-null    object 
 7   UNIT_MULT: Unit of Multiplier    630 non-null    object 
 8   OBS_STATUS: Observation Status   131 non-null    object 
 9   OBS_COMMENT: Observation Comment  0 non-null      float64
dtypes: float64(2), object(8)
memory usage: 49.3+ KB
```

We can see there are no missing value in the other data. The missing data in OBS_STATUS: ObservationStatus and OBS_COMMENT: Observation Comment would not affect much to the dataset since all of the other necessary data like Measure, Region, OBS_Value been provided without being missed in any columns.

Below we check if any data is null.

```
df.isnull()
```

	DATAFLOW	MEASURE: Measure	REGION: Region	FREQ: Frequency	TIME_PERIOD: Time Period	OBS_VALUE	UNIT_MEASURE: Unit o
0	False	False	False	False	False	False	
1	False	False	False	False	False	False	
2	False	False	False	False	False	False	
3	False	False	False	False	False	False	
4	False	False	False	False	False	False	
...	...	...	...	...	...	...	
625	False	False	False	False	False	False	
626	False	False	False	False	False	False	
627	False	False	False	False	False	False	
628	False	False	False	False	False	False	
629	False	False	False	False	False	False	

630 rows × 10 columns

It returns true since there are a lot of null value in both OBS_STATUS and OBS_COMMENT, but there is no missing value on other ones (Region, Frequency, etc), which shows as return false.

```
df.isnull().sum()
```

```
DATAFLOW                           0
MEASURE: Measure                   0
REGION: Region                     0
FREQ: Frequency                    0
TIME_PERIOD: Time Period           0
OBS_VALUE                          0
UNIT_MEASURE: Unit of Measure      0
UNIT_MULT: Unit of Multiplier      0
OBS_STATUS: Observation Status   499
```

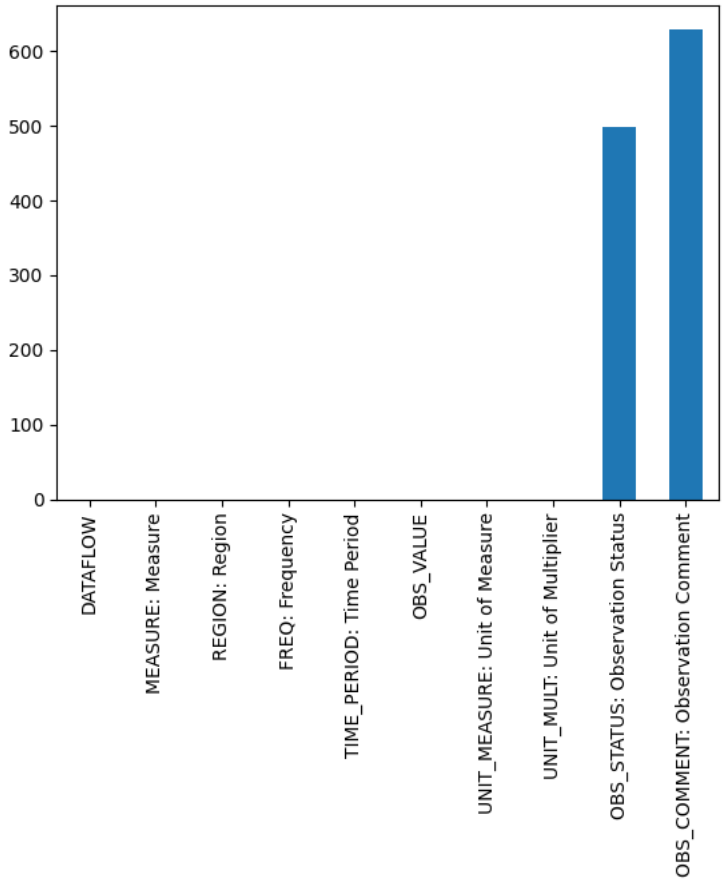
```
OBS_COMMENT: Observation Comment    630
dtype: int64
```

There are a lot of null data in OBS_STATUS: Observation Status and OBS_COMMENT: Observation Comment. Let's go to the data further.

```
df.isnull().any(axis=1)
```

```
0      True
1      True
2      True
3      True
4      True
       ...
625    True
626    True
627    True
628    True
629    True
Length: 630, dtype: bool
```

```
df.isnull().sum().plot.bar()
plt.show()
```




```
df[df.isnull().any(axis=1)]
```

	DATAFLOW	MEASURE: Measure	REGION: Region	FREQ: Frequency	TIME_PERIOD: Time Peri
0	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-(
1	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-(
2	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-(
3	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	2020-(

```
df.duplicated()
```

```
0      False
1      False
2      False
3      False
4      False
       ...
625    False
626    False
627    False
628    False
629    False
Length: 630, dtype: bool
```

There are no duplicated values.

```
df.describe()
```

	OBS_VALUE	OBS_COMMENT: Observation Comment
count	6.300000e+02	0.0
mean	7.980742e+05	NaN
std	1.948605e+06	NaN
min	8.570000e+01	NaN
25%	9.138750e+02	NaN
50%	2.967640e+04	NaN
75%	4.139467e+05	NaN
max	1.009407e+07	NaN

We now create 2 variable called categorical and numerical to make it easier for inspecting the columns given their different characteristics.

```
# List of categorical variables
categorical = [i for i in df.columns if df[i].dtypes == 'O']
# List of numerical variables
numerical = [i for i in df.columns if i not in categorical]
print('categorical:', categorical, '\n', 'numerical: ', numerical)
```

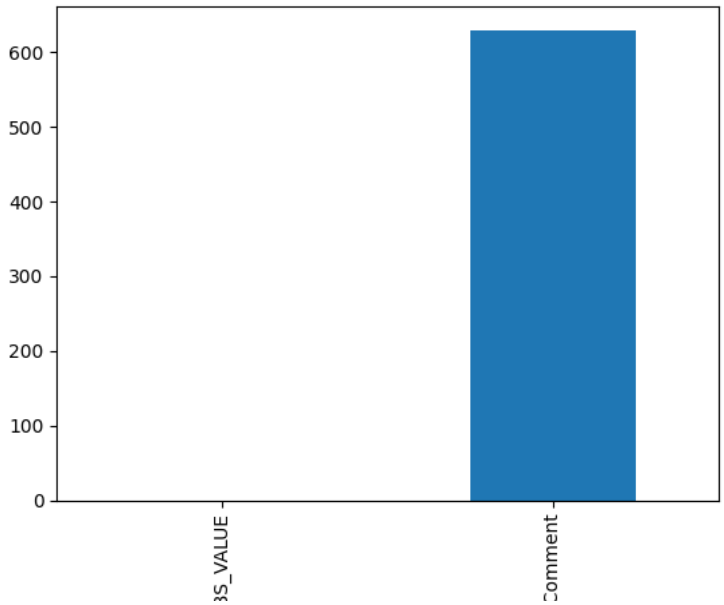
```
categorical: ['DATAFLOW', 'MEASURE: Measure', 'REGION: Region', 'FREQ: Frequency', 'TIME_PERIOD: Time Period', 'UNIT_MEA
 numerical:  ['OBS_VALUE', 'OBS_COMMENT: Observation Comment']
```

Let's check how many null values are there in each variable

```
df[numerical].isnull().sum()
```

```
OBS_VALUE                          0
OBS_COMMENT: Observation Comment    630
dtype: int64
```

```
df[numerical].isnull().sum().plot.bar()
plt.show()
```



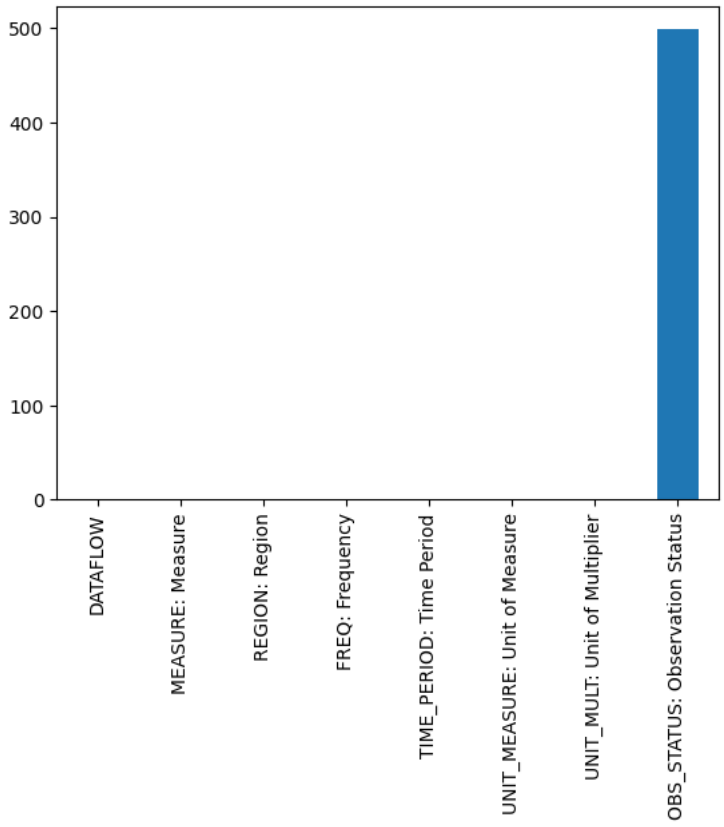
As can be seen, OBS_COMMENT: Observation Comment is the only one missing values in numerical segment.

Now let's check the null value in the categorical section.

```
df[categorical].isnull().sum()
```

```
DATAFLOW                          0
MEASURE: Measure                  0
REGION: Region                    0
FREQ: Frequency                   0
TIME_PERIOD: Time Period          0
UNIT_MEASURE: Unit of Measure     0
UNIT_MULT: Unit of Multiplier     0
OBS_STATUS: Observation Status    499
dtype: int64
```

```
df[categorical].isnull().sum().plot.bar()
plt.show()
```



As can be seen, OBS_COMMENT: Observation Comment is the only one missing values in categorical segment.

```
df[categorical].describe()
```

	DATAFLOW	MEASURE: Measure	REGION: Region	FREQ: Frequency	TIME_PERIOD: Time Pe
count	630	630	630	630	
unique	1	5	9	1	
top	ABS:RES_DWELL_ST(1.0.0)	2: Value of dwelling stock: Owned by households	4: South Australia	Q: Quarterly	20
freq	630	126	70	630	

From the summary table, we can see the most frequent values of each variable and its frequency in the dataset.

2. Housing Price and Number of Residential Units From 2020-Q1 To 2023-Q2

In this section, analyzing the mean dwelling price and the number of dwelling units will be provided. From these statistics, it would reflect the healthy of housing market value, which could be considered as a source for buyers/sellers to understand the housing market's health and trends.

First of all, let's have a look at the housing price on 2020 at quarter 1.

```
def housing_price_covid_time(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2023-Q2
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]

    return f"The dwelling price in {region} in 2020-Q1 is {filtered_df}."


#Input the data to calculate the 'mean price of residential dwelling' in each state
resultAUS = housing_price_covid_time(df, 'AUS: Australia', '5: Mean price of residential dwellings')
resultNSW = housing_price_covid_time(df, '1: New South Wales', '5: Mean price of residential dwellings')
resultVIC = housing_price_covid_time(df, '2: Victoria', '5: Mean price of residential dwellings')
resultQLD = housing_price_covid_time(df, '3: Queensland', '5: Mean price of residential dwellings')
resultSA = housing_price_covid_time(df, '4: South Australia', '5: Mean price of residential dwellings')
resultWA = housing_price_covid_time(df, '5: Western Australia', '5: Mean price of residential dwellings')
resultTAS = housing_price_covid_time(df, '6: Tasmania', '5: Mean price of residential dwellings')
resultNT = housing_price_covid_time(df, '7: Northern Territory', '5: Mean price of residential dwellings')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The dwelling price in AUS: Australia in 2020-Q1 is 694.7.
The dwelling price in 1: New South Wales in 2020-Q1 is 895.8.
The dwelling price in 2: Victoria in 2020-Q1 is 761.2.
The dwelling price in 3: Queensland in 2020-Q1 is 524.9.
The dwelling price in 4: South Australia in 2020-Q1 is 465.3.
The dwelling price in 5: Western Australia in 2020-Q1 is 501.3.
The dwelling price in 6: Tasmania in 2020-Q1 is 460.1.
The dwelling price in 7: Northern Territory in 2020-Q1 is 415.9.
```

The dwelling price in New South Wales was highest in Australia being $895.8. It could be understood and proves that Sydney is the biggest city in Australia and the housing market there is competitive.

In contrast, the housing price in Northern Territory was the lowest one being only $415.9.

Now we want to find out the current mean dwelling price (2023-Q2) in each state

```
def current_housing_price(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2023-Q2
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]

    return f"The current dwelling price in {region} in 2023-Q2 is {filtered_df}."


#Input the data to calculate the 'mean price of residential dwelling' in each state
resultAUS = current_housing_price(df, 'AUS: Australia', '5: Mean price of residential dwellings')
```

```
resultNSW = current_housing_price(df, '1: New South Wales', '5: Mean price of residential dwellings')
resultVIC = current_housing_price(df, '2: Victoria', '5: Mean price of residential dwellings')
resultQLD = current_housing_price(df, '3: Queensland', '5: Mean price of residential dwellings')
resultSA = current_housing_price(df, '4: South Australia', '5: Mean price of residential dwellings')
resultWA = current_housing_price(df, '5: Western Australia', '5: Mean price of residential dwellings')
resultTAS = current_housing_price(df, '6: Tasmania', '5: Mean price of residential dwellings')
resultNT = current_housing_price(df, '7: Northern Territory', '5: Mean price of residential dwellings')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The current dwelling price in AUS: Australia in 2023-Q2 is 912.7.
The current dwelling price in 1: New South Wales in 2023-Q2 is 1167.5.
The current dwelling price in 2: Victoria in 2023-Q2 is 904.8.
The current dwelling price in 3: Queensland in 2023-Q2 is 781.6.
The current dwelling price in 4: South Australia in 2023-Q2 is 684.7.
The current dwelling price in 5: Western Australia in 2023-Q2 is 671.0.
The current dwelling price in 6: Tasmania in 2023-Q2 is 662.2.
The current dwelling price in 7: Northern Territory in 2023-Q2 is 521.7.
```

From the data above, we can see that the highest dwelling price is New South Wales being $1167.5.

In constrast, the lowest one currently is Northern Territory being $521.7.

It can also be seen that the housing price difference between New South Wales and Northern have been escalated more since during the last three years, the dwelling rise in New South Wales is higher comparing to the one in Northern Territory.

Now let's find out the number of individual residential units like homes, condos, apartment, etc in each state, so we can get an idea of housing supply to compare between the house supply from 2020 to 2023. This metric with the mean price statistics above could be used to analyze the housing market from 2020 up to now.

```
def calculate_percentage_change(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to include only the relevant quarters (2020-Q1 and 2023-Q2)
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]

    # Check if data for the specified quarters and region exists
    if len(filtered_df) != 2:
        return f"Insufficient data for {region} and measure '{measure}' for 2020-Q1 and 2023-Q2"

    # Get the average prices for 2020-Q1 and 2023-Q2
    price_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
    price_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]

    # Calculate the percentage change
    percentage_change = ((price_2023_q2 - price_2020_q1) / price_2020_q1) * 100

    return f"The percentage change in the {measure} in {region} from 2020-Q1 to 2023-Q2 is {percentage_change:.2f}%."


# Calculate the 'mean price of residential dwelling' in each state
resultAUS = calculate_percentage_change(df, 'AUS: Australia', '5: Mean price of residential dwellings')
resultNSW = calculate_percentage_change(df, '1: New South Wales', '5: Mean price of residential dwellings')
resultVIC = calculate_percentage_change(df, '2: Victoria', '5: Mean price of residential dwellings')
resultQLD = calculate_percentage_change(df, '3: Queensland', '5: Mean price of residential dwellings')
resultSA = calculate_percentage_change(df, '4: South Australia', '5: Mean price of residential dwellings')
resultWA = calculate_percentage_change(df, '5: Western Australia', '5: Mean price of residential dwellings')
resultTAS = calculate_percentage_change(df, '6: Tasmania', '5: Mean price of residential dwellings')
resultNT = calculate_percentage_change(df, '7: Northern Territory', '5: Mean price of residential dwellings')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The percentage change in the 5: Mean price of residential dwellings in AUS: Australia from 2020-Q1 to 2023-Q2 is 31.38%.
The percentage change in the 5: Mean price of residential dwellings in 1: New South Wales from 2020-Q1 to 2023-Q2 is 30.
The percentage change in the 5: Mean price of residential dwellings in 2: Victoria from 2020-Q1 to 2023-Q2 is 18.86%.
The percentage change in the 5: Mean price of residential dwellings in 3: Queensland from 2020-Q1 to 2023-Q2 is 48.90%.
The percentage change in the 5: Mean price of residential dwellings in 4: South Australia from 2020-Q1 to 2023-Q2 is 47.
The percentage change in the 5: Mean price of residential dwellings in 5: Western Australia from 2020-Q1 to 2023-Q2 is 3
```

```
The percentage change in the 5: Mean price of residential dwellings in 6: Tasmania from 2020-Q1 to 2023-Q2 is 43.93%.
The percentage change in the 5: Mean price of residential dwellings in 7: Northern Territory from 2020-Q1 to 2023-Q2 is
```

Now, let's analyze the data further by seeing which states have the higher housing mean price increase comparing to the Australian average mean price. has the highest percentage change and which one is the lowest one.

```
results = {
        'AUS': 31.88,
        'NSW': 30.33,
        'VIC': 18.86,
        'QLD': 48.90,
        'SA': 47.15,
        'WA': 33.85,
        'TAS': 43.93,
        'NT': 25.44
    }

def compare_with_aus_price():

    # Initialize lists to hold regions that are bigger and smaller than AUS
    bigger_than_aus = []
    smaller_than_aus = []

    # Extract the value for AUS
    result_aus = results.get('AUS')

    # Compare each region with AUS
    for region, value in results.items():
        if value > result_aus:
            bigger_than_aus.append(region)
        else:
            smaller_than_aus.append(region)

    return bigger_than_aus, smaller_than_aus

bigger, smaller = compare_with_aus_price()
print(f"Regions with percentage change bigger than AUS: {bigger}")
print(f"Regions with percentage change smaller than AUS: {smaller}")
```

```
Regions with percentage change bigger than AUS: ['QLD', 'SA', 'WA', 'TAS']
Regions with percentage change smaller than AUS: ['AUS', 'NSW', 'VIC', 'NT']
```

Let's check which state has the highest percentage change and which one is the lowest one.

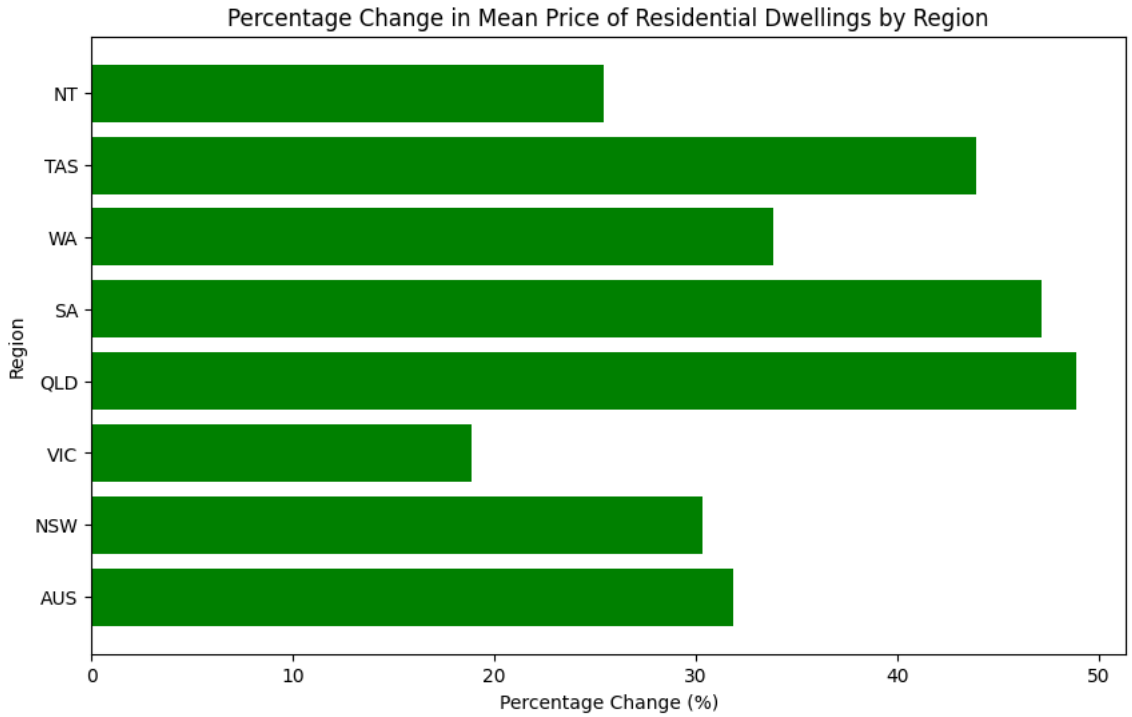
```
# Initialize variables to hold the highest and smallest values and their corresponding keys
highest_percentage_value = None
highest_percentage_key = None
smallest_percentage_value = None
smallest_percentage_key = None

# Iterate through the dictionary to find the highest and smallest values
for key, value in results.items():
    if highest_percentage_value is None or value > highest_percentage_value:
        highest_percentage_value = value
        highest_percentage_key = key

    if smallest_percentage_value is None or value < smallest_percentage_value:
        smallest_percentage_value = value
        smallest_percentage_key = key

# Output the results
print(f"The highest percentage change is: {highest_percentage_key} with {highest_percentage_value}%")
print(f"The smallest percentage change is: {smallest_percentage_key} with {smallest_percentage_value}%")
```

```
The highest percentage change is: QLD with 48.9%
The smallest percentage change is: VIC with 18.86%
```

We will now visualise by using bar chart, so it would be easier to analyze the percentage change in mean price of residential dwellings from 2020 to 2023.

```
def draw_histogram(results):
    keys = list(results.keys())
    values = list(results.values())

    plt.figure(figsize=(10, 6))
    plt.barh(keys, values, color='green')
    plt.xlabel('Percentage Change (%)')
```

```
    plt.ylabel('Region')
    plt.title('Percentage Change in Mean Price of Residential Dwellings by Region')

    plt.show()

draw_histogram(results)
```



As we can see from the data above, the highest increase regarding the dwelling mean price in the last three years is Queensland, which proves that the housing demand in here is much higher. It could imply that more and more people have moved to Queensland recently. It is followed by South Australia being 47.15% and Tasmania being 43.93%.

It could be seen that the increase in housing mean price in Victoria is the lowest one being only 18.86%, which could reflect that the housing demand in Victoria is not high and there might be a migration from Victoria to other states more than the immigration (We might need to look at other datatset relating to the immigration/migration to get the conclusion though).

3. Number of Residential Dwellings In Each State

In this section, we will find out which state has the highest and lowest figure for residential units in both 2020 and 2023. Moreover, we will analyze the percentage change in the residential unit supply during the last three years.

From these data combining with the mean dwelling price in the above section, it could provide the report to analyze the housing market's health and trends.

```
def no_residential_dwelling(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2020-Q1
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]

    return f"The number of individual residential units in {region} in 2020-Q1 is {filtered_df}."


#Input the data to calculate the 'mean price of residential dwelling' in each state
resultAUS = no_residential_dwelling(df, 'AUS: Australia', '4: Number of residential dwellings')
resultNSW = no_residential_dwelling(df, '1: New South Wales', '4: Number of residential dwellings')
resultVIC = no_residential_dwelling(df, '2: Victoria', '4: Number of residential dwellings')
resultQLD = no_residential_dwelling(df, '3: Queensland', '4: Number of residential dwellings')
resultSA = no_residential_dwelling(df, '4: South Australia', '4: Number of residential dwellings')
resultWA = no_residential_dwelling(df, '5: Western Australia', '4: Number of residential dwellings')
resultTAS = no_residential_dwelling(df, '6: Tasmania', '4: Number of residential dwellings')
resultNT = no_residential_dwelling(df, '7: Northern Territory', '4: Number of residential dwellings')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
```

```
print(resultNT);
```

```
The number of individual residential units in AUS: Australia in 2020-Q1 is 10476.6.
The number of individual residential units in 1: New South Wales in 2020-Q1 is 3246.3.
The number of individual residential units in 2: Victoria in 2020-Q1 is 2717.6.
The number of individual residential units in 3: Queensland in 2020-Q1 is 2098.0.
The number of individual residential units in 4: South Australia in 2020-Q1 is 789.9.
The number of individual residential units in 5: Western Australia in 2020-Q1 is 1112.9.
The number of individual residential units in 6: Tasmania in 2020-Q1 is 247.7.
The number of individual residential units in 7: Northern Territory in 2020-Q1 is 85.7.
```

From the data above, we can see that the highest number of individual residential units in 2020 quarter 1 was New South Wales being 3,246.3 while the smallest one was Northern Territory only being 85.7.

Let's check the figure for residential dwellings in 2023 quarter 2.

```
def no_residential_dwelling(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2023-Q2
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]

    return f"The number of individual residential units in {region} in 2023-Q2 is {filtered_df}."


#Input the data to calculate the 'mean price of residential dwelling' in each state
resultAUS = no_residential_dwelling(df, 'AUS: Australia', '4: Number of residential dwellings')
resultNSW = no_residential_dwelling(df, '1: New South Wales', '4: Number of residential dwellings')
resultVIC = no_residential_dwelling(df, '2: Victoria', '4: Number of residential dwellings')
resultQLD = no_residential_dwelling(df, '3: Queensland', '4: Number of residential dwellings')
resultSA = no_residential_dwelling(df, '4: South Australia', '4: Number of residential dwellings')
resultWA = no_residential_dwelling(df, '5: Western Australia', '4: Number of residential dwellings')
resultTAS = no_residential_dwelling(df, '6: Tasmania', '4: Number of residential dwellings')
resultNT = no_residential_dwelling(df, '7: Northern Territory', '4: Number of residential dwellings')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The number of individual residential units in AUS: Australia in 2023-Q2 is 11055.8.
The number of individual residential units in 1: New South Wales in 2023-Q2 is 3425.6.
The number of individual residential units in 2: Victoria in 2023-Q2 is 2893.4.
The number of individual residential units in 3: Queensland in 2023-Q2 is 2220.6.
The number of individual residential units in 4: South Australia in 2023-Q2 is 817.8.
The number of individual residential units in 5: Western Australia in 2023-Q2 is 1151.3.
The number of individual residential units in 6: Tasmania in 2023-Q2 is 262.3.
The number of individual residential units in 7: Northern Territory in 2023-Q2 is 89.3.
```

The figure has been increased in all of the states from 2020 quarter 1 to 2023 quarter 2 and the order is still the same as previously with New South Wales remains at the top place of house supply (3,425.6) while Northern Territory was still at the bottom.

Let's calculate how much each state change from 2020 quarter to 2023 quarter regarding the number of house supply.

```
def calc_percentage_change(df, region, measure):
  # Filter to get the specific data for region and measure include rows for the specified region and measure
  filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

  # Filter the data to include only the relevant quarters (2020-Q1 and 2023-Q2)
  filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]

  no_dwelling_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
  no_dwelling_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]
  percentage_change = ((no_dwelling_2023_q2 - no_dwelling_2020_q1) / no_dwelling_2020_q1) * 100

  return f"The percentage change in the {measure} in {region} from 2020-Q1 to 2023-Q2 is {percentage_change:.2f}%."

resultAUS = calc_percentage_change(df, 'AUS: Australia', '4: Number of residential dwellings')
resultNSW = calc_percentage_change(df, '1: New South Wales', '4: Number of residential dwellings')
resultVIC = calc_percentage_change(df, '2: Victoria', '4: Number of residential dwellings')
resultQLD = calc_percentage_change(df, '3: Queensland', '4: Number of residential dwellings')
resultSA = calc_percentage_change(df, '4: South Australia', '4: Number of residential dwellings')
resultWA = calc_percentage_change(df, '5: Western Australia', '4: Number of residential dwellings')
resultTAS = calc_percentage_change(df, '6: Tasmania', '4: Number of residential dwellings')
resultNT = calc_percentage_change(df, '7: Northern Territory', '4: Number of residential dwellings')
print(resultAUS);
```

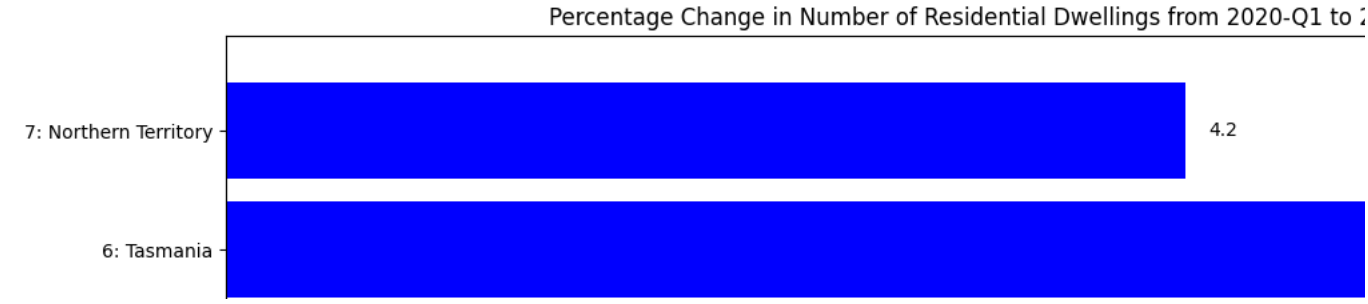
```
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The percentage change in the 4: Number of residential dwellings in AUS: Australia from 2020-Q1 to 2023-Q2 is 5.53%.
The percentage change in the 4: Number of residential dwellings in 1: New South Wales from 2020-Q1 to 2023-Q2 is 5.52%.
The percentage change in the 4: Number of residential dwellings in 2: Victoria from 2020-Q1 to 2023-Q2 is 6.47%.
The percentage change in the 4: Number of residential dwellings in 3: Queensland from 2020-Q1 to 2023-Q2 is 5.84%.
The percentage change in the 4: Number of residential dwellings in 4: South Australia from 2020-Q1 to 2023-Q2 is 3.53%.
The percentage change in the 4: Number of residential dwellings in 5: Western Australia from 2020-Q1 to 2023-Q2 is 3.45%
The percentage change in the 4: Number of residential dwellings in 6: Tasmania from 2020-Q1 to 2023-Q2 is 5.89%.
The percentage change in the 4: Number of residential dwellings in 7: Northern Territory from 2020-Q1 to 2023-Q2 is 4.20
```

As we can see from the above result, the state that has the highest percentage of residential dwelling change from 2020 to 2023 is Victoria being 6.47% increase, while the lowest one is Western Australia is Western Australia with only 3.45%.

Now let's use a bar chart to see the data clearer.

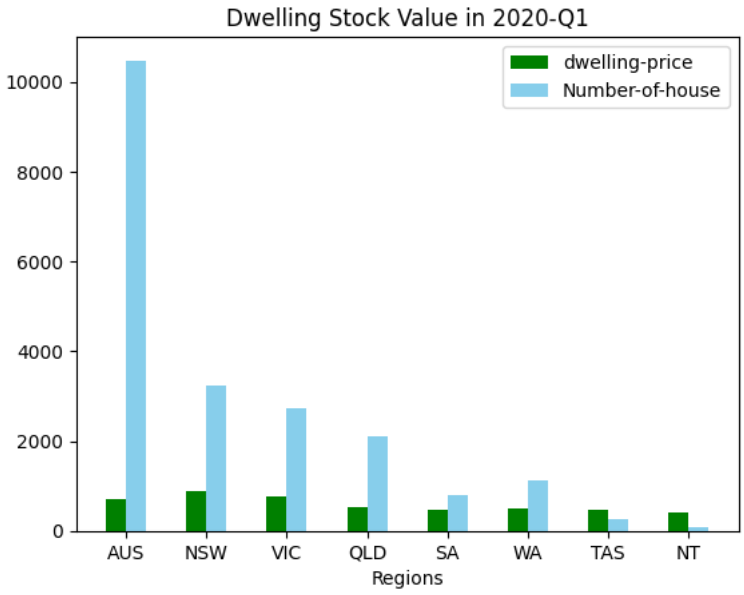
```
#Update the cal_percentage_change function above
def calc_percentage_change(df, region, measure):
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]
    no_dwelling_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
    no_dwelling_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]
    percentage_change = ((no_dwelling_2023_q2 - no_dwelling_2020_q1) / no_dwelling_2020_q1) * 100
    return percentage_change

# Calculate percentage changes
regions = ['AUS: Australia', '1: New South Wales', '2: Victoria', '3: Queensland', '4: South Australia', '5: Western Australia
percentage_changes = {region: calc_percentage_change(df, region, '4: Number of residential dwellings') for region in regions}

# Prepare key and values for plotting
labels = list(percentage_changes.keys())
values = list(percentage_changes.values())

# Create the bar chart
plt.figure(figsize=(15, 10))
plt.barh(labels, values, color='blue')
plt.xlabel('Percentage Change (%)')
plt.title('Percentage Change in Number of Residential Dwellings from 2020-Q1 to 2023-Q2')
for i, v in enumerate(values):
    plt.text(v + 0.1, i, str(round(v, 2)), color='black', verticalalignment='center')
plt.show()
```



Based on the data above, we can see that the housing market has been recovered in all of the states. As you can see that the housing supply has been risen up, with the highest increase is Victoria (6.47%) followed by Tasmania (5.89%). However, Perth would need to focus on finding out ways to improve house supply (for instance, government invest more funds if it's possible to offer more houses for residents) since the individual residential units in Western Australia only increasing 3.45%.

We will also combine the dataset of both the mean dwelling price and the number of residential dwelling in each state between 2020 quarter 1 and 2023 quarter 2. Firstly, let's look at the data in 2020 firstly.

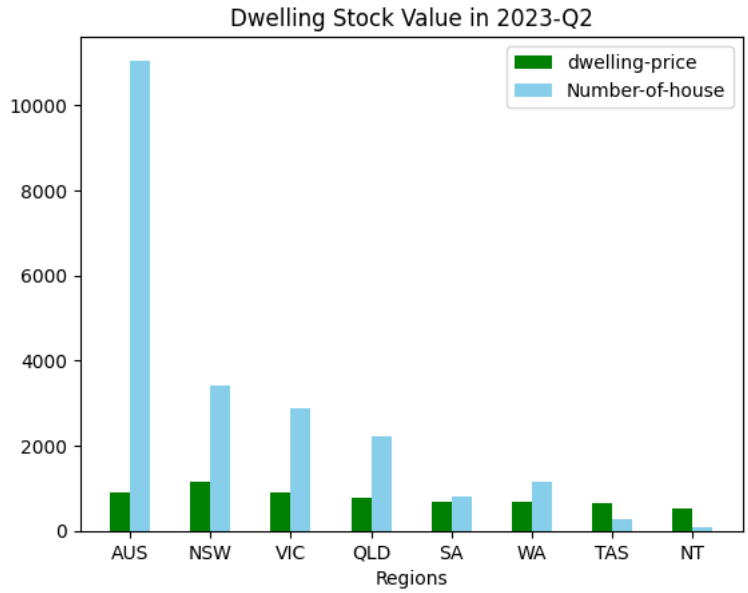
```
regions = ['AUS', 'NSW', 'VIC', 'QLD', 'SA', 'WA', 'TAS', 'NT']
mean_dwelling_price = [694.7, 895.8, 761.2, 524.9, 465.3, 501.3, 460.1, 415.9]
no_dwelling = [10476.6, 3246.3, 2717.6, 2098.0, 789.9, 1112.9, 247.7, 85.7]

x = np.arange(len(regions))
width = 0.25

plt.bar(x - width/2, mean_dwelling_price, width, label='dwelling-price', color='green')
plt.bar(x + width/2, no_dwelling, width, label='Number-of-house', color='skyblue')

plt.xlabel('Regions')
plt.title('Dwelling Stock Value in 2020-Q1')
plt.xticks(x, regions)
plt.legend()

plt.show()
```



Now let's look at the data in 2023 quarter 2.

```
regions = ['AUS', 'NSW', 'VIC', 'QLD', 'SA', 'WA', 'TAS', 'NT']
mean_dwelling_price = [912.7, 1167.5, 904.8, 781.6, 684.7, 671.0, 662.2, 521.7]
no_dwelling = [11055.8, 3425.6, 2893.4, 2220.6, 817.8, 1151.3, 262.3, 89.3]

x = np.arange(len(regions))
width = 0.25

plt.bar(x - width/2, mean_dwelling_price, width, label='dwelling-price', color='green')
plt.bar(x + width/2, no_dwelling, width, label='Number-of-house', color='skyblue')

plt.xlabel('Regions')
plt.title('Dwelling Stock Value in 2023-Q2')
plt.xticks(x, regions)
plt.legend()
```

```
plt.show()
```



As we can see from the figure above, from 2020 - Q1 to 2023 - Q2, the order of the highest to the lowest housing price and number of house supply have not been changed with NSW being at the highest in both sectors during the last three years, while Northern Territory remains at the lowest dwelling price and house supply.

Although NSW has the highest housing price, the state still also have the highest house supply, It can imply that both the supply and demand are high as although there are a lot of residential dwelling, the price keeps increase after these three years, proving that the demand is also high.

In contrast, the house supply in Northern Territory is low and the demand is not high as well, reflecting by the statistics it still remains at the bottom in both housing price being $521.7 and the number of residential units being 89.7 (only rise 4.2% more from 2020-Q1).

4. Conclusion From The Result of The Housing Mean Price & The Number of Residential Units

We can see from the data in the two above sections that New South Wales is still at top for both house supply and the mean dwelling price in the last three years. There mean price is still increase up to 30.33% between 2020 quarter 1 and 2023 quarter 2, meaning the buyer's demand is still immensive there. It is suggested that the government might need to invest further regarding the house supply (increase 5.52% which is still lower than Victoria and Queensland) if it's possible so as to give more opportunities for buyers, which will also help to balance the power between the buyers and sellers.

We can also see from the above data & statistic that the number of residential units in Victoria has increased 6.47%, which is the highest figure. Although Victoria's housing price is still high ($904.8), the mean price has increased only 18.86% in the last three years, which is lower than many other states. It could be implied that the housing market in the last three years in Victoria has offered more advantages for buyers since housing supply has been increased and the price has not been escalated as other states.

Tasmania has grown quick from 2020 to currently as well because the mean housing price has been increased significantly to 43.93% and the house supply has grown being 5.89%. The data shows us that there might be an abundance of immigrants coming to Tasmania, which rise the demand to get houses there.

It could also be seen that Queensland's housing price has been increasing significantly to 48.90% only in the last three years, and the housing supply has also be increased up to 5.84%, just below Victoria and Tasmania. It could show that the number of immigrants to the state have been increased and the state is more developed in the last few years.

Last but not least, from the report, it could be shown that Northern Territory had the lowest growth in both housing mean price and number of house supply, which shows the housing market demand there is not high and there has been not much development regarding to house market there in the last few years.

5. Dwelling Stock Value Owned By Household and Non Household

We will now go to the next section to compare the dwelling stock value owned by household and non household like corporations, investment funds, and government entities. Additionally, we will check the housing stock value recovery after Covid by checking the percentage change from quarter 1, 2020 to currently quarter 2, 2023.

```
def dwelling_stock_household(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2020-Q1
```

```
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]

    return f"The dwelling stock owned by household in {region} in 2020-Q1 is {filtered_df}."


#Input the data to calculate the residential dwelling stock which are owned by household in each state
resultAUS = dwelling_stock_household(df, 'AUS: Australia', '2: Value of dwelling stock: Owned by households')
resultNSW = dwelling_stock_household(df, '1: New South Wales', '2: Value of dwelling stock: Owned by households')
resultVIC = dwelling_stock_household(df, '2: Victoria', '2: Value of dwelling stock: Owned by households')
resultQLD = dwelling_stock_household(df, '3: Queensland', '2: Value of dwelling stock: Owned by households')
resultSA = dwelling_stock_household(df, '4: South Australia', '2: Value of dwelling stock: Owned by households')
resultWA = dwelling_stock_household(df, '5: Western Australia', '2: Value of dwelling stock: Owned by households')
resultTAS = dwelling_stock_household(df, '6: Tasmania', '2: Value of dwelling stock: Owned by households')
resultNT = dwelling_stock_household(df, '7: Northern Territory', '2: Value of dwelling stock: Owned by households')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The dwelling stock owned by household in AUS: Australia in 2020-Q1 is 6949644.6.
The dwelling stock owned by household in 1: New South Wales in 2020-Q1 is 2766544.7.
The dwelling stock owned by household in 2: Victoria in 2020-Q1 is 2008574.3.
The dwelling stock owned by household in 3: Queensland in 2020-Q1 is 1049595.8.
The dwelling stock owned by household in 4: South Australia in 2020-Q1 is 344736.8.
The dwelling stock owned by household in 5: Western Australia in 2020-Q1 is 529575.8.
The dwelling stock owned by household in 6: Tasmania in 2020-Q1 is 106974.4.
The dwelling stock owned by household in 7: Northern Territory in 2020-Q1 is 27911.1.
```

From the result above, we could see that in 2020 the total market value of all residential properties that are owned by individual households are $6,949,644.6. The highest market value was New South Wales with being 2,766544.7 while the lowest one was Northern Territory.

Now let's display it visually with the chart below.

```
def dwelling_stock_household(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2020-Q1
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]

    return f"The dwelling stock owned by household in {region} in 2023-Q2 is {filtered_df}."


#Input the data to calculate the residential dwelling stock which are owned by household in each state
resultAUS = dwelling_stock_household(df, 'AUS: Australia', '2: Value of dwelling stock: Owned by households')
resultNSW = dwelling_stock_household(df, '1: New South Wales', '2: Value of dwelling stock: Owned by households')
resultVIC = dwelling_stock_household(df, '2: Victoria', '2: Value of dwelling stock: Owned by households')
resultQLD = dwelling_stock_household(df, '3: Queensland', '2: Value of dwelling stock: Owned by households')
resultSA = dwelling_stock_household(df, '4: South Australia', '2: Value of dwelling stock: Owned by households')
resultWA = dwelling_stock_household(df, '5: Western Australia', '2: Value of dwelling stock: Owned by households')
resultTAS = dwelling_stock_household(df, '6: Tasmania', '2: Value of dwelling stock: Owned by households')
resultNT = dwelling_stock_household(df, '7: Northern Territory', '2: Value of dwelling stock: Owned by households')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The dwelling stock owned by household in AUS: Australia in 2023-Q2 is 9694790.4.
The dwelling stock owned by household in 1: New South Wales in 2023-Q2 is 3838849.5.
The dwelling stock owned by household in 2: Victoria in 2023-Q2 is 2546676.3.
The dwelling stock owned by household in 3: Queensland in 2023-Q2 is 1665041.1.
The dwelling stock owned by household in 4: South Australia in 2023-Q2 is 530888.6.
The dwelling stock owned by household in 5: Western Australia in 2023-Q2 is 738219.8.
The dwelling stock owned by household in 6: Tasmania in 2023-Q2 is 164289.5.
The dwelling stock owned by household in 7: Northern Territory in 2023-Q2 is 37804.8.
```

We can see that the figure of dwelling stock possessed by household has increased significantly to 9,694,790.4 with the highest one still being New South Wales ($3,838849.5) and the lowest one is still Northern Territory being 37,804.8.

Now let's check the percentage of change from 2020 to 2023 in each state and also Australia in general.

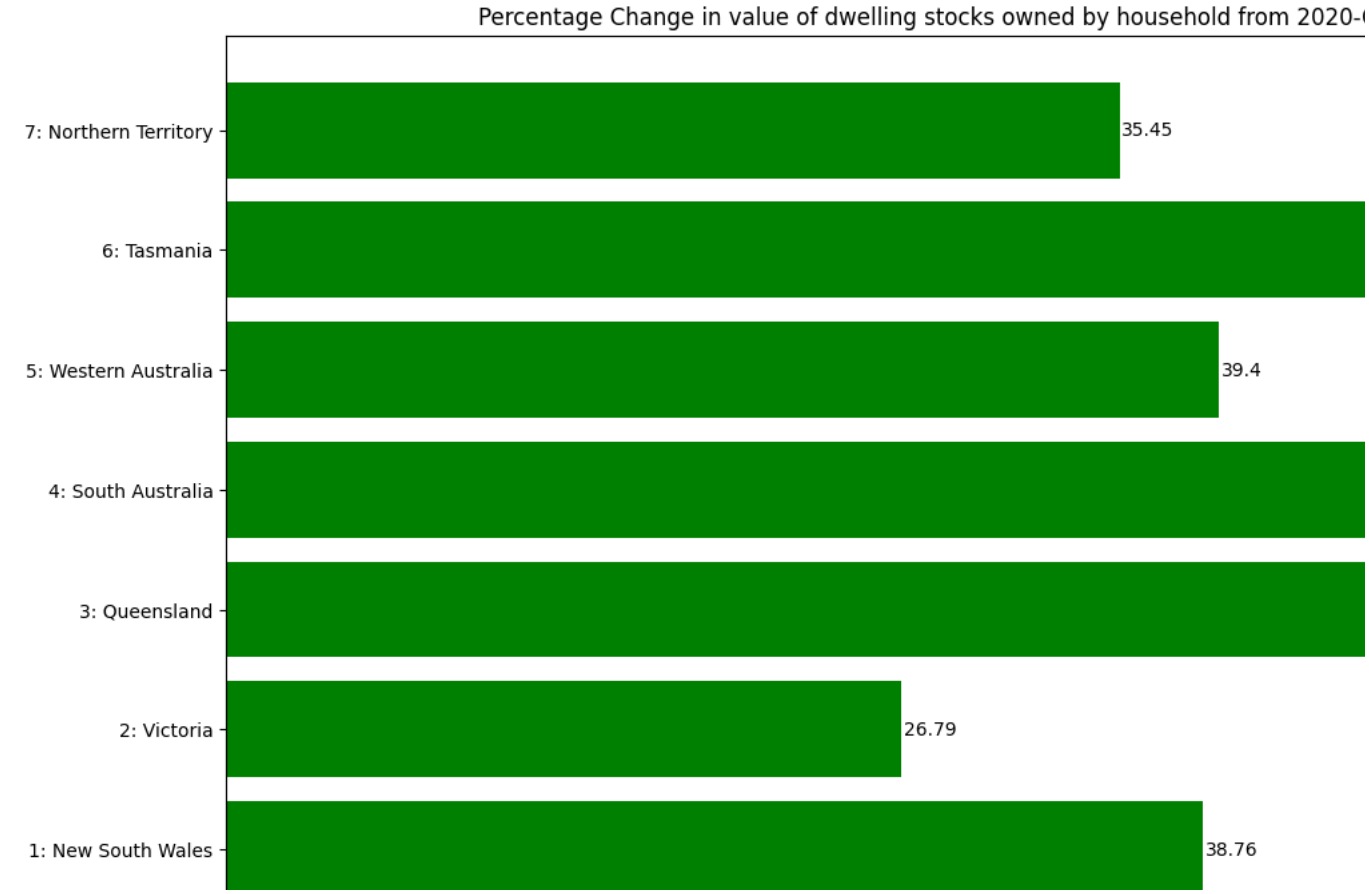
```
def calc_percentage_change(df, region, measure):
  # Filter to get the specific data for region and measure include rows for the specified region and measure
  filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

  # Filter the data to include only the relevant quarters (2020-Q1 and 2023-Q2)
  filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]

  no_dwelling_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
  no_dwelling_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]
  percentage_change = ((no_dwelling_2023_q2 - no_dwelling_2020_q1) / no_dwelling_2020_q1) * 100

  return f"The percentage change in the {measure} in {region} from 2020-Q1 to 2023-Q2 is {percentage_change:.2f}%."

resultAUS = calc_percentage_change(df, 'AUS: Australia', '2: Value of dwelling stock: Owned by households')
resultNSW = calc_percentage_change(df, '1: New South Wales', '2: Value of dwelling stock: Owned by households')
resultVIC = calc_percentage_change(df, '2: Victoria', '2: Value of dwelling stock: Owned by households')
resultQLD = calc_percentage_change(df, '3: Queensland', '2: Value of dwelling stock: Owned by households')
resultSA = calc_percentage_change(df, '4: South Australia', '2: Value of dwelling stock: Owned by households')
resultWA = calc_percentage_change(df, '5: Western Australia', '2: Value of dwelling stock: Owned by households')
resultTAS = calc_percentage_change(df, '6: Tasmania', '2: Value of dwelling stock: Owned by households')
resultNT = calc_percentage_change(df, '7: Northern Territory', '2: Value of dwelling stock: Owned by households')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The percentage change in the 2: Value of dwelling stock: Owned by households in AUS: Australia from 2020-Q1 to 2023-Q2 i
The percentage change in the 2: Value of dwelling stock: Owned by households in 1: New South Wales from 2020-Q1 to 2023-
The percentage change in the 2: Value of dwelling stock: Owned by households in 2: Victoria from 2020-Q1 to 2023-Q2 is 2
The percentage change in the 2: Value of dwelling stock: Owned by households in 3: Queensland from 2020-Q1 to 2023-Q2 is
The percentage change in the 2: Value of dwelling stock: Owned by households in 4: South Australia from 2020-Q1 to 2023-
The percentage change in the 2: Value of dwelling stock: Owned by households in 5: Western Australia from 2020-Q1 to 202
The percentage change in the 2: Value of dwelling stock: Owned by households in 6: Tasmania from 2020-Q1 to 2023-Q2 is 5
The percentage change in the 2: Value of dwelling stock: Owned by households in 7: Northern Territory from 2020-Q1 to 20
```

From the calculation above, we could see that from 2020 up to the quarter 2 of 2023 the highest increase regarding the value of dwelling stock owned by households is Queensland with being 58.64%, following by South Australia being 54%. In contrast, the lowest increase is Victoria with only being 26.79%.

Now let's show it clearer from the chart displayed below.

```
#Update the cal_percentage_change function above
def calc_percentage_change(df, region, measure):
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]
    no_dwelling_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
    no_dwelling_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]
    percentage_change = ((no_dwelling_2023_q2 - no_dwelling_2020_q1) / no_dwelling_2020_q1) * 100
    return percentage_change

# Calculate percentage changes
regions = ['AUS: Australia', '1: New South Wales', '2: Victoria', '3: Queensland', '4: South Australia', '5: Western Australia
percentage_changes = {region: calc_percentage_change(df, region, '2: Value of dwelling stock: Owned by households') for region

# Prepare key and values for plotting
labels = list(percentage_changes.keys())
values = list(percentage_changes.values())

# Create the bar chart
plt.figure(figsize=(15, 10))
plt.barh(labels, values, color='green')
plt.xlabel('Percentage Change (%)')
plt.title('Percentage Change in value of dwelling stocks owned by household from 2020-Q1 to 2023-Q2')
for i, v in enumerate(values):
    plt.text(v + 0.1, i, str(round(v, 2)), color='black', verticalalignment='center')
plt.show()
```



Now we will check the dwelling stock owned by non household including corporations, investment funds, and government entities. From this, we can compare the stock value owned between household and non-household.

Firstly, let's have a look at the value of dwelling stock owned by non household in 2020-quarter 1.

```
def dwelling_stock_nonhousehold(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2020-Q1
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]

    return f"The dwelling stock owned by non household in {region} in 2020-Q1 is {filtered_df}."


#Input the data to calculate the residential dwelling stock which are owned by non household in each state
resultAUS = dwelling_stock_nonhousehold(df, 'AUS: Australia', '3: Value of dwelling stock: Owned by non households')
resultNSW = dwelling_stock_nonhousehold(df, '1: New South Wales', '3: Value of dwelling stock: Owned by non households')
resultVIC = dwelling_stock_nonhousehold(df, '2: Victoria', '3: Value of dwelling stock: Owned by non households')
resultQLD = dwelling_stock_nonhousehold(df, '3: Queensland', '3: Value of dwelling stock: Owned by non households')
resultSA = dwelling_stock_nonhousehold(df, '4: South Australia', '3: Value of dwelling stock: Owned by non households')
resultWA = dwelling_stock_nonhousehold(df, '5: Western Australia', '3: Value of dwelling stock: Owned by non households')
resultTAS = dwelling_stock_nonhousehold(df, '6: Tasmania', '3: Value of dwelling stock: Owned by non households')
resultNT = dwelling_stock_nonhousehold(df, '7: Northern Territory', '3: Value of dwelling stock: Owned by non households')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

```
The dwelling stock owned by non household in AUS: Australia in 2020-Q1 is 328428.1.
The dwelling stock owned by non household in 1: New South Wales in 2020-Q1 is 141417.7.
The dwelling stock owned by non household in 2: Victoria in 2020-Q1 is 60135.8.
The dwelling stock owned by non household in 3: Queensland in 2020-Q1 is 51659.2.
The dwelling stock owned by non household in 4: South Australia in 2020-Q1 is 22847.8.
The dwelling stock owned by non household in 5: Western Australia in 2020-Q1 is 28310.7.
The dwelling stock owned by non household in 6: Tasmania in 2020-Q1 is 6993.7.
The dwelling stock owned by non household in 7: Northern Territory in 2020-Q1 is 7734.0.
```

As we can observe from the data above, the highest dwelling stock value owned by non-household was in New South Wales being $141,417.7.

The lowest one was Tasmania with only being $6993.7

```
def dwelling_stock_nonhousehold(df, region, measure):
    # Filter to get the specific data for region and measure include rows for the specified region and measure
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]

    # Filter the data to see the current dwelling price in 2020-Q1
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]

    return f"The dwelling stock owned by non household in {region} in 2023-Q2 is {filtered_df}."


#Input the data to calculate the residential dwelling stock which are owned by non household in each state
resultAUS = dwelling_stock_nonhousehold(df, 'AUS: Australia', '3: Value of dwelling stock: Owned by non households')
resultNSW = dwelling_stock_nonhousehold(df, '1: New South Wales', '3: Value of dwelling stock: Owned by non households')
resultVIC = dwelling_stock_nonhousehold(df, '2: Victoria', '3: Value of dwelling stock: Owned by non households')
resultQLD = dwelling_stock_nonhousehold(df, '3: Queensland', '3: Value of dwelling stock: Owned by non households')
resultSA = dwelling_stock_nonhousehold(df, '4: South Australia', '3: Value of dwelling stock: Owned by non households')
resultWA = dwelling_stock_nonhousehold(df, '5: Western Australia', '3: Value of dwelling stock: Owned by non households')
resultTAS = dwelling_stock_nonhousehold(df, '6: Tasmania', '3: Value of dwelling stock: Owned by non households')
resultNT = dwelling_stock_nonhousehold(df, '7: Northern Territory', '3: Value of dwelling stock: Owned by non households')
print(resultAUS);
print(resultNSW);
print(resultVIC);
print(resultQLD);
print(resultSA);
print(resultWA);
print(resultTAS);
print(resultNT);
```

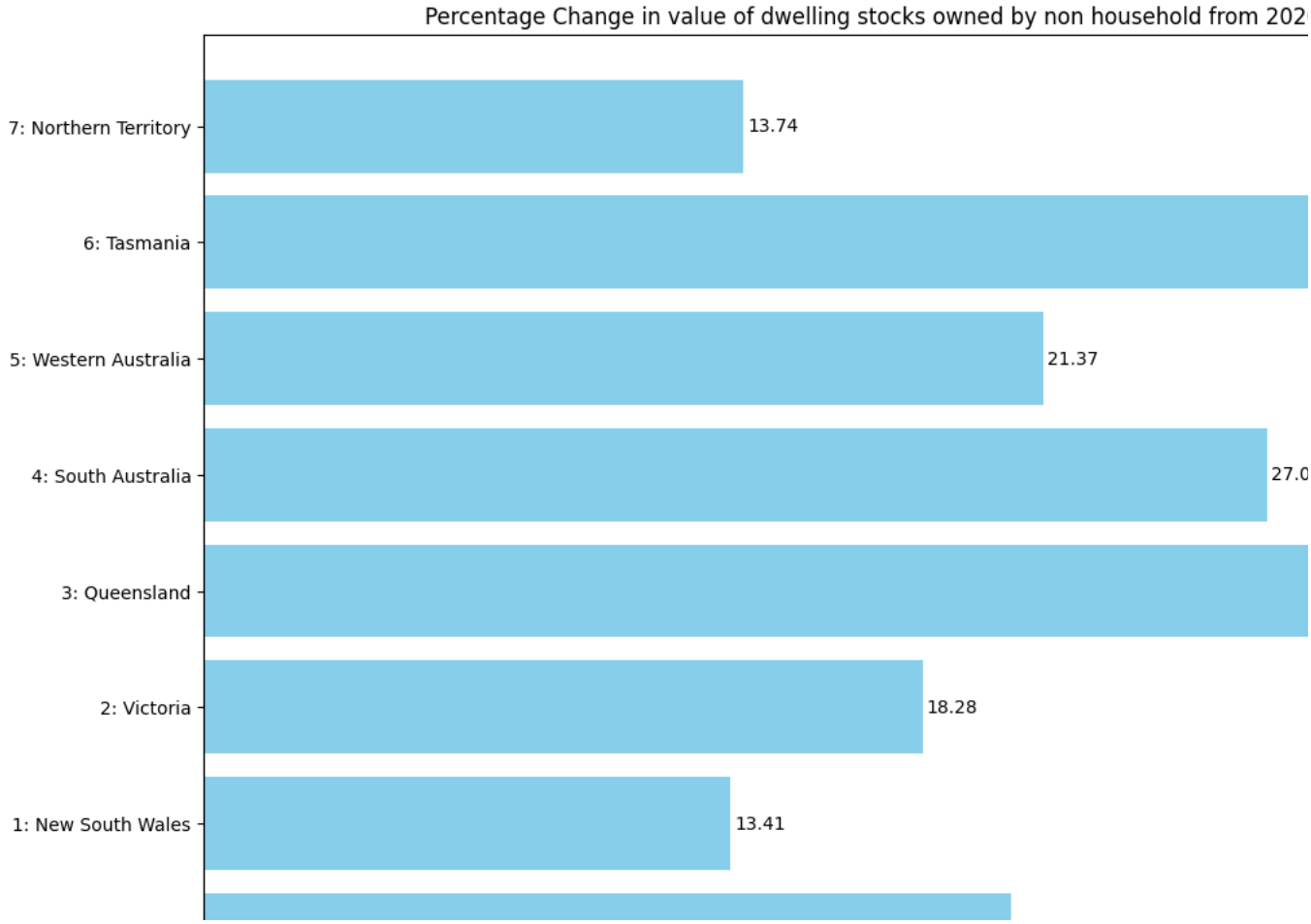
```
The dwelling stock owned by non household in AUS: Australia in 2023-Q2 is 395833.9.
The dwelling stock owned by non household in 1: New South Wales in 2023-Q2 is 160375.7.
The dwelling stock owned by non household in 2: Victoria in 2023-Q2 is 71131.4.
The dwelling stock owned by non household in 3: Queensland in 2023-Q2 is 70491.5.
The dwelling stock owned by non household in 4: South Australia in 2023-Q2 is 29027.6.
The dwelling stock owned by non household in 5: Western Australia in 2023-Q2 is 34361.3.
The dwelling stock owned by non household in 6: Tasmania in 2023-Q2 is 9371.5.
The dwelling stock owned by non household in 7: Northern Territory in 2023-Q2 is 8796.6.
```

As we can see from above, the highest dwelling stock owned by non-household is New South Wales being $160,375.7.

In contrast, the smallest one is Northern Territory with only $8,796.6

Now let's see the percentage change of dwelling stock owned by non-household from quarter 1, 2020 to quarter 2, 2023.

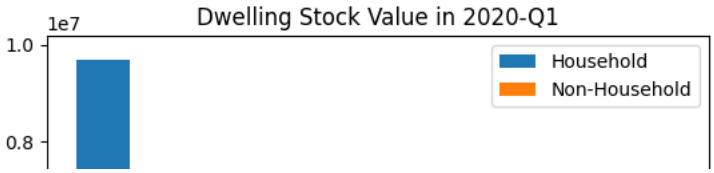
```
#Update the cal_percentage_change function above
def calc_percentage_change(df, region, measure):
    filtered_df = df[(df['REGION: Region'] == region) & (df['MEASURE: Measure'] == measure)]
    filtered_df = filtered_df[filtered_df['TIME_PERIOD: Time Period'].isin(['2020-Q1', '2023-Q2'])]
    no_dwelling_2020_q1 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2020-Q1']['OBS_VALUE'].values[0]
    no_dwelling_2023_q2 = filtered_df[filtered_df['TIME_PERIOD: Time Period'] == '2023-Q2']['OBS_VALUE'].values[0]
    percentage_change = ((no_dwelling_2023_q2 - no_dwelling_2020_q1) / no_dwelling_2020_q1) * 100
    return percentage_change

# Calculate percentage changes
regions = ['AUS: Australia', '1: New South Wales', '2: Victoria', '3: Queensland', '4: South Australia', '5: Western Australia
percentage_changes = {region: calc_percentage_change(df, region, '3: Value of dwelling stock: Owned by non households') for re

# Prepare key and values for plotting
labels = list(percentage_changes.keys())
values = list(percentage_changes.values())

# Create the bar chart
plt.figure(figsize=(15, 10))
plt.barh(labels, values, color='skyblue')
plt.xlabel('Percentage Change (%)')
plt.title('Percentage Change in value of dwelling stocks owned by non household from 2020-Q1 to 2023-Q2')
for i, v in enumerate(values):
    plt.text(v + 0.1, i, str(round(v, 2)), color='black', verticalalignment='center')
plt.show()
```



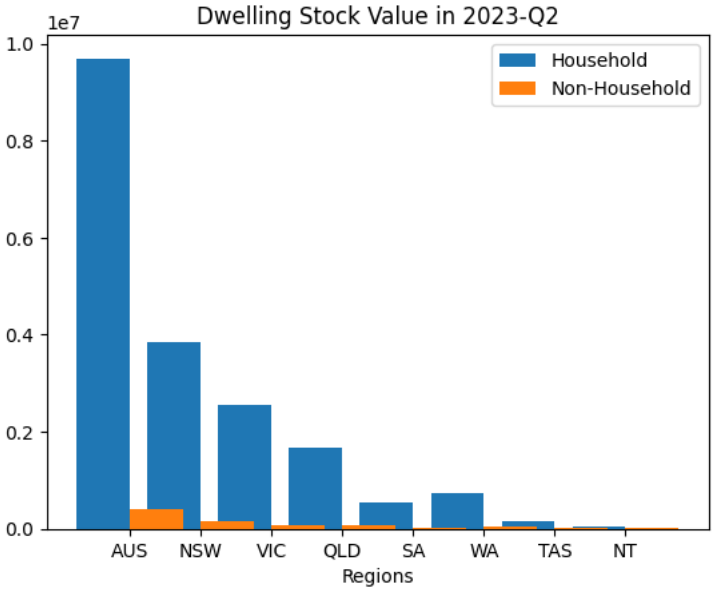
As we can see, the highest percentage change of dwelling stock owned by non-household sector is Queensland being 36.45% while the lowest one is Victoria with only being 13.41% and it is followed by Tasmania (34%).

Now let's compare the value of dwelling stock owned by household and non-household (corporations, investment funds, and government entities) in both 2020 and 2023.

```
regions = ['AUS', 'NSW', 'VIC', 'QLD', 'SA', 'WA', 'TAS', 'NT']
household_values_2020Q1 = [6949644.6, 2766544.7, 2008574.3, 1049595.8, 344736.8, 529575.8, 106974.4, 27911.1]
non_household_values_2020Q1 = [328428.1, 141417.7, 60135.8, 51659.2, 22847.8, 28310.7, 6993.7, 7734.0]

# X-axis locations

x = np.arange(len(regions))
width = 0.75

plt.bar(x - width/2, household_values, width, label='Household')
plt.bar(x + width/2, non_household_values, width, label='Non-Household')

plt.xlabel('Regions')
plt.title('Dwelling Stock Value in 2020-Q1')
plt.xticks(x, regions)
plt.legend()

plt.show()
```




```
regions = ['AUS', 'NSW', 'VIC', 'QLD', 'SA', 'WA', 'TAS', 'NT']
household_values = [9694790.4, 3838849.5, 2546676.3, 1665041.1, 530888.6, 738219.8, 164289.5, 37804.8]
non_household_values = [395833.9, 160375.7, 71131.4, 70491.5, 29027.6, 34361.3, 9371.5, 8796.6]

x = np.arange(len(regions))
width = 0.75

plt.bar(x - width/2, household_values, width, label='Household')
plt.bar(x + width/2, non_household_values, width, label='Non-Household')

plt.xlabel('Regions')
plt.title('Dwelling Stock Value in 2023-Q2')
plt.xticks(x, regions)
plt.legend()

plt.show()
```



From the data above regarding dwelling stock value owned by household and non-household in both 2020 and 2023, we can see that most of the stock owned by household in all of the states and it totally dominates the non household sector.

6. Summary

In the nutshell, we can see that New South Wales still has the highest figure in both house supply as well as the mean dwelling price in both 2020 and 2023. The market in New South Wales is running well with high demand and high supply based on the statistics calculated above and it has kept recovery well after the COVID. Another point that could be pointed out is that Queensland's dwelling price has been significantly increased to 48.90% in only three years, which reflect the increase demand to purchase in this state. Melbourne has shown the lowest increment in housing price comparing to the other states, being only 18.86%, although the house supply has enhanced the most being 6.47%, which shows that currently the house demand has been dropped there currently. In contrast, Northern Territory remains the lowest growth in both housing mean price and number of house supply, which shows the housing market demand and supply there are not high and has not been improved much since 2020.

The other report shows that the value of dwelling stock owned by household has been dominant the ones owned by non household between 2020 and 2023. During these three years, the dwelling stock owned by non-household has been escalated drastically in Queensland being 36.45% rise following by Tasmania being 34.0%. The same figure has been shown for the household sector as Queensland has experienced the boost of 58.64% for the dwelling stock value following by Tasmania being 53.58%.