

Singapore Housing System (2021–2025)

This section explores Singapore's public housing market using official HDB data published by the Ministry of National Development (MND). It combines two datasets: one containing resale flat transactions and the other containing flat rental records. The aim is to analyze trends in resale prices and rental yields across towns and flat types from January 2021 to May 2025.

Table of contents:

1. Introduction
2. Executive Summary
3. Data Description & Cleaning
4. Analysis and Key Findings
5. Conclusions
6. Further Analysis
7. References

1. Introduction

Since its founding in 1960, the Housing & Development Board (HDB) has been central to Singapore's public housing, evolving from solving slum conditions to becoming a socio-economic benchmark.

Official records show that over 80% of Singapore's population lives in HDB flats spread across 24 towns and three estates, with more than one million flats completed to date.

This study focuses on the post-COVID period from January 2021 to May 2025—to assess market affordability. It combines two official datasets (resale transactions and flat rental records) to compare median resale prices and rental yields across towns and flat types. This analysis uncovers insights into how rental returns measure against resale values in the public housing landscape.

Research questions:

This study aims to explore the following questions:

1. Is there a strong correlation between resale price per sqm and monthly rent?

2. Which towns offer the highest rent-to-price (rental yield) ratios?
3. How does flat age relate to both resale and rental prices?
4. Have rental rates grown faster than resale prices over the past five years?
5. Do mature and non-mature towns show clear differences in average resale prices and rental yields?
6. How do resale prices and rental yields for 3-room flats in Sengkang compare over time?

Data Files:

This report utilizes two datasets—resale flat transactions and flat rental records, and linked by the fields town and flat_type. Due to their size, the original CSV files are provided in compressed (zipped) format.

RentingFlats.zip

The rental dataset was downloaded from the Singapore Government's open data portal:

[Renting Out of Flats 2025 –](#)

[HDB&resultId=d_c9f57187485a850908655db0e8cfe651&page=1\)](#)

It was published by the Housing and Development Board (HDB). The dataset includes the following fields:

1. Rental Approval Date
2. Town
3. Block
4. Street name
5. Flat type
6. Monthly rent

ResaleFlats.zip

The resale dataset was downloaded from [Resale flat prices based on registration date from Jan-2017 onwards](#).

It was published by the Housing and Development Board (HDB). The dataset includes the following fields:

1. Month
2. Town
3. Flat Type
4. Block

5. Street Name
6. Storey Range
7. Floor Area Sqm
8. Flat Model
9. Lease Commence Date
10. Remaining Lease
11. Resale Price

2. Executive Summary

When analyzing HDB rental data, it is important to consider occupancy regulations, which cap the number of occupants allowed per flat type.

For example:

Flat Type	Maximum Occupants
1 & 2 room	4 occupants
3-room	6 occupants
4-room	8 occupants

In addition to above rules, rental yield is affected by other variables such as:

1. Flat Type
2. Town (Area)
3. Flat Size

Understanding these constraints and variables provides valuable context for analyzing differences in median rent, resale prices, and rental yields across various HDB flat types and towns

3.Data Description & Cleaning

```
In [207... # Import Pandas, Matplotlib and seaborn modules
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [208... # Load ResaleFlats and RentingFlats zipped datasets
resale_df = pd.read_csv("ResaleFlats.zip")
rent_df = pd.read_csv("RentingFlats.zip")

# Convert 'month' column to datetime format
resale_df['month'] = pd.to_datetime(resale_df['month'], format='%Y

# Filter for records from Jan 2021 onward
```

```
resale_df_filtered = resale_df[resale_df['month'] >= '2021-01-01']

# Understand the datasets
print(resale_df_filtered.columns.is_unique)
resale_df.isnull().sum()
print(resale_df_filtered.shape)
print(resale_df_filtered[resale_df.duplicated() == True])
resale_df_filtered.head()
```

True

(121336, 11)

	month	town	flat_type	block	street_name	st
orey_range \						
89038	2021-01-01	PUNGGOL	4 ROOM	260C	PUNGGOL WAY	
13 TO 15						
94486	2021-03-01	WOODLANDS	4 ROOM	889D	WOODLANDS DR	50
07 TO 09						
95478	2021-04-01	HOUGANG	3 ROOM	473B	UPP SERANGOON CRES	
04 TO 06						
96024	2021-04-01	PUNGGOL	4 ROOM	308A	PUNGGOL WALK	
01 TO 03						
96105	2021-04-01	PUNGGOL	5 ROOM	663B	PUNGGOL DR	
16 TO 18						
...
...						
204170	2025-06-01	SEMPAWANG	2 ROOM	103A	CANBERRA ST	
04 TO 06						
204411	2025-04-01	SEMPAWANG	4 ROOM	103B	CANBERRA ST	
04 TO 06						
204413	2025-04-01	SEMPAWANG	4 ROOM	103B	CANBERRA ST	
04 TO 06						
205294	2025-04-01	SENGKANG	4 ROOM	457B	SENGKANG WEST RD	
13 TO 15						
208433	2025-05-01	YISHUN	4 ROOM	426B	YISHUN AVE	11
10 TO 12						

	floor_area_sqm	flat_model	lease_commence_date	\
89038	93.0	Model A	2016	
94486	93.0	Model A	2017	
95478	68.0	Model A	2017	
96024	92.0	Premium Apartment	2016	
96105	110.0	Improved	2016	
...
204170	38.0	2-room	2020	
204411	93.0	Model A	2020	
204413	93.0	Model A	2020	
205294	93.0	Model A	2020	
208433	92.0	Model A	2015	

	remaining_lease	resale_price
89038	94 years 06 months	490000.0
94486	95 years 05 months	475000.0
95478	95 years 02 months	360000.0
96024	94 years	500000.0
96105	94 years 03 months	650000.0
...
204170	94 years 07 months	350000.0
204411	94 years 09 months	610000.0
204413	94 years 09 months	610000.0
205294	94 years 06 months	685000.0
208433	88 years 11 months	595000.0

[85 rows x 11 columns]

```
/tmp/ipykernel_111/2034957623.py:15: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
```

```
print(resale_df_filtered[resale_df.duplicated() == True])
```

Out[208...

	month	town	flat_type	block	street_name	storey_range	floor_ar
87589	2021-01-01	ANG MO KIO	2 ROOM	170	ANG MO KIO AVE 4	01 TO 03	
87590	2021-01-01	ANG MO KIO	2 ROOM	170	ANG MO KIO AVE 4	07 TO 09	
87591	2021-01-01	ANG MO KIO	3 ROOM	331	ANG MO KIO AVE 1	04 TO 06	
87592	2021-01-01	ANG MO KIO	3 ROOM	534	ANG MO KIO AVE 10	04 TO 06	
87593	2021-01-01	ANG MO KIO	3 ROOM	561	ANG MO KIO AVE 10	01 TO 03	

Since this study focuses on resale data from January 2021 onward, the month column was converted to datetime format to enable filtering. After filtering, we confirmed that the dataset contains only records from 2021 onward. The shape of the filtered dataset was examined, and duplicate records were checked. The dataset appears to contain unique entries.

In [209...

```
# Clean the resale data
# Convert resale price from string to numeric
resale_df_filtered['resale_price'] = pd.to_numeric(resale_df_filt

# Convert floor_area_sqm from string to numeric
resale_df_filtered['floor_area_sqm'] = pd.to_numeric(resale_df_fil

# Convert month to datetime format and add the new column txn_date
resale_df_filtered['txn_date'] = pd.to_datetime(resale_df_filtered

# Remove the null or invalid values from the columns
resale_df_filtered = resale_df_filtered.drop_duplicates(
    subset=['month', 'town', 'flat_type', 'floor_area_sqm']
)

# Calculate price per square meter and add as new column
resale_df_filtered['price_per_sqm'] = resale_df_filtered['resale_p

# Extract year and month from 'txn_date'
resale_df_filtered['year'] = resale_df_filtered['txn_date'].dt.yea
resale_df_filtered['month'] = resale_df_filtered['txn_date'].dt.mo

resale_df_filtered.info()
# Grouping by year and calculating average values
```

```
yearly_stats = resale_df_filtered.groupby('year')[['resale_price',  
print(yearly_stats)
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
Index: 42790 entries, 87589 to 208924
```

```
Data columns (total 14 columns):
```

#	Column	Non-Null Count	Dtype
0	month	42790 non-null	int32
1	town	42790 non-null	object
2	flat_type	42790 non-null	object
3	block	42790 non-null	object
4	street_name	42790 non-null	object
5	storey_range	42790 non-null	object
6	floor_area_sqm	42790 non-null	float64
7	flat_model	42790 non-null	object
8	lease_commence_date	42790 non-null	int64
9	remaining_lease	42790 non-null	object
10	resale_price	42790 non-null	float64
11	txn_date	42790 non-null	datetime64[ns]
12	price_per_sqm	42790 non-null	float64
13	year	42790 non-null	int32

```
dtypes: datetime64[ns](1), float64(3), int32(2), int64(1), object(7)
```

```
memory usage: 4.6+ MB
```

	year	resale_price	price_per_sqm
0	2021	516070.320304	5057.677216
1	2022	555668.600186	5507.904065
2	2023	580749.097338	5859.854486
3	2024	644177.965705	6490.267173
4	2025	677226.919449	6920.556822

```
/tmp/ipykernel_111/520300752.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
resale_df_filtered['resale_price'] = pd.to_numeric(resale_df_filtered['resale_price'])
```

```
/tmp/ipykernel_111/520300752.py:6: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
resale_df_filtered['floor_area_sqm'] = pd.to_numeric(resale_df_filtered['floor_area_sqm'])
```

```
/tmp/ipykernel_111/520300752.py:9: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
resale_df_filtered['txn_date'] = pd.to_datetime(resale_df_filtered['month'], format='%Y-%m')
```

Resale prices have shown a steady increase over the analyzed period, suggesting a positive market trend.

The dataset is grouped by town, flat type, year, and month. Median values for resale price and price per square meter are computed, along with transaction counts.

```
In [210... # Group data monthly  
resale_monthly = resale_df_filtered.groupby(['town', 'flat_type',  
      median_resale_price=('resale_price', 'median'),  
      median_price_per_sqm=('price_per_sqm', 'median'),  
      resale_txn_count=('txn_date', 'count')  
)  
  
print(resale_monthly)
```


	town	flat_type	year	month	median_resale_pric
e \					
0	ANG MO KIO	2 ROOM	2021	1	211000.
0					
1	ANG MO KIO	2 ROOM	2021	4	213250.
0					
2	ANG MO KIO	2 ROOM	2021	5	230000.
0					
3	ANG MO KIO	2 ROOM	2021	7	230000.
0					
4	ANG MO KIO	2 ROOM	2021	9	233000.
0					
...	
...					
6215	YISHUN	MULTI-GENERATION	2023	12	975000.
0					
6216	YISHUN	MULTI-GENERATION	2024	2	998000.
0					
6217	YISHUN	MULTI-GENERATION	2024	3	1200000.
0					
6218	YISHUN	MULTI-GENERATION	2024	6	1028000.
0					
6219	YISHUN	MULTI-GENERATION	2025	5	945000.
0					
	median_price_per_sqm	resale_txn_count			
0	4688.888889	1			
1	4793.939394	2			
2	5227.272727	1			
3	5167.929293	2			
4	5295.454545	1			
...			
6215	5446.927374	1			
6216	6085.365854	1			
6217	7317.073171	1			
6218	6011.695906	1			
6219	6428.571429	1			

[6220 rows x 7 columns]

```
In [211... # Understand the rent datasets
print(rent_df.columns.is_unique)
rent_df.isnull().sum()
print(rent_df.shape)
print(rent_df[rent_df.duplicated() == True])
rent_df.head()
```

```

True
(162130, 6)
      rent_approval_date      town block      street_name fl
at_type \
243      2021-01      BEDOK      99      BEDOK NTH AVE 4
5-R00M
305      2021-01      BEDOK      525      BEDOK NTH ST 3
3-R00M
510      2021-01      BUKIT BATOK      341      BT BATOK ST 34
4-R00M
591      2021-01      BUKIT MERAH      17      TELOK BLANGAH CRES
5-R00M
672      2021-01      BUKIT MERAH      120      BT MERAH VIEW
4-R00M
...      ...      ...      ...      ...
...
155863      2025-04      HOUGANG      305      HOUGANG AVE 5
3-R00M
156220      2025-04      BISHAN      455      SIN MING AVE
4-R00M
157315      2025-04      SENGKANG      320B      ANCHORVALE DR
5-R00M
159531      2025-05      PASIR RIS      628      PASIR RIS DR 3
5-R00M
160181      2025-05      QUEENSTOWN      83      C'WEALTH CL
3-R00M

```

```

      monthly_rent
243      2000
305      1600
510      2000
591      2550
672      2400
...      ...
155863      2900
156220      2800
157315      3900
159531      2000
160181      2650

```

```
[279 rows x 6 columns]
```

```

/tmp/ipykernel_111/3250356717.py:5: UserWarning: Boolean Series key
will be reindexed to match DataFrame index.
  print(rent_df[resale_df.duplicated() == True])

```

Out[211...

	rent_approval_date	town	block	street_name	flat_type	monthly_rent
0	2021-01	ANG MO KIO	105	ANG MO KIO AVE 4	4-ROOM	2000
1	2021-01	ANG MO KIO	107	ANG MO KIO AVE 4	3-ROOM	1750
2	2021-01	ANG MO KIO	108	ANG MO KIO AVE 4	3-ROOM	1750
3	2021-01	ANG MO KIO	111	ANG MO KIO AVE 4	5-ROOM	2230
4	2021-01	ANG MO KIO	111	ANG MO KIO AVE 4	5-ROOM	2450

The shape of the rental dataset has been examined, comprising 162,130 records across 6 columns. No duplicate entries were found, ensuring the data is clean and suitable for trend analysis.

In [212...

```
# Clean the rent data
# Convert 'monthly_rent' column from string to numeric
rent_df['monthly_rent'] = pd.to_numeric(rent_df['monthly_rent'])

# Convert 'rent_approval_date' to datetime format and store as 'rent_date'
rent_df['rent_date'] = pd.to_datetime(rent_df['rent_approval_date'])

# Remove the null or invalid values from the columns
rent_df = rent_df.drop_duplicates(
    subset=['monthly_rent', 'rent_date']
)

# Extract year and month from 'rent_date'
rent_df['year'] = rent_df['rent_date'].dt.year
rent_df['month'] = rent_df['rent_date'].dt.month

rent_df.info()
# Grouping by year and calculating average values
yearly_stats = rent_df.groupby('year')[['monthly_rent']].mean().reset_index()
print(yearly_stats)
```

```

<class 'pandas.core.frame.DataFrame'>
Index: 5734 entries, 0 to 161804
Data columns (total 9 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   rent_approval_date                    5734 non-null   object
1   town                                  5734 non-null   object
2   block                                 5734 non-null   object
3   street_name                           5734 non-null   object
4   flat_type                             5734 non-null   object
5   monthly_rent                          5734 non-null   int64
6   rent_date                             5734 non-null   datetime64[ns]
7   year                                  5734 non-null   int32
8   month                                 5734 non-null   int32
dtypes: datetime64[ns](1), int32(2), int64(1), object(5)
memory usage: 403.2+ KB
   year  monthly_rent
0  2021    2134.853594
1  2022    2587.374587
2  2023    3014.428571
3  2024    3022.434438
4  2025    3105.160410

```

The average monthly rent has shown a consistent upward trend from 2021 to 2025, reflecting a steady increase in rental demand or pricing over the post-COVID recovery period.

The dataset is grouped by town, flat type, year, and month. Median value for monthly rent is computed, along with block counts.

```

In [213...] rent_monthly = rent_df.groupby(['town', 'flat_type', 'year', 'month'],
      median_rent=('monthly_rent', 'median'),
      rent_txn_count=('block', 'count'))
print(rent_monthly)

```

	town	flat_type	year	month	median_rent	rent_txn_count
0	ANG MO KIO	2-ROOM	2021	1	1400.0	
2						
1	ANG MO KIO	2-ROOM	2021	2	1500.0	
2						
2	ANG MO KIO	2-ROOM	2021	3	1025.0	
2						
3	ANG MO KIO	2-ROOM	2021	4	950.0	
2						
4	ANG MO KIO	2-ROOM	2021	5	1000.0	
1						
...	
...						
2652	YISHUN	EXECUTIVE	2024	2	3189.0	
2						
2653	YISHUN	EXECUTIVE	2024	3	5500.0	
1						
2654	YISHUN	EXECUTIVE	2024	9	3848.0	
1						
2655	YISHUN	EXECUTIVE	2024	12	3588.0	
1						
2656	YISHUN	EXECUTIVE	2025	2	5600.0	
1						

[2657 rows x 6 columns]

Before merging, we standardized the 'flat_type' column across both datasets to avoid mismatch caused by inconsistent formatting.

```
In [214... # Normalize flat_type in both datasets
resale_monthly['flat_type'] = resale_monthly['flat_type'].astype(str)
resale_df_filtered['flat_type'] = resale_df_filtered['flat_type'].value_counts()
```

```
Out[214... flat_type
4 ROOM      15268
5 ROOM      11436
3 ROOM       9163
EXECUTIVE    5309
2 ROOM       1546
MULTI-GENERATION  41
1 ROOM        27
Name: count, dtype: int64
```

```
In [215... rent_monthly['flat_type'] = rent_monthly['flat_type'].astype(str)
rent_df['flat_type'] = rent_df['flat_type'].value_counts()
```

```
Out[215... flat_type
3-ROOM      1997
4-ROOM      1789
5-ROOM      1320
2-ROOM       340
EXECUTIVE    278
1-ROOM       10
Name: count, dtype: int64
```

```
In [216... # Perform the merge
merged_df = pd.merge(
    resale_monthly,
    rent_monthly,
    on=['town', 'flat_type', 'year', 'month'],
    how='inner'
)
```

```
In [217... # To calculate the rental rental_yield
# We would need to convert Monthly rent into annual rent
merged_df['annual_rent'] = merged_df['median_rent'] * 12
# To calculate the rental yield, we would need to divide annual re
merged_df['rental_yield_pct'] = merged_df['annual_rent'] / merged_

#Convert to upper case and removes empty spaces
merged_df['town'] = merged_df['town'].astype(str).str.upper().str.
merged_df['flat_type'] = merged_df['flat_type'].astype(str).str.up

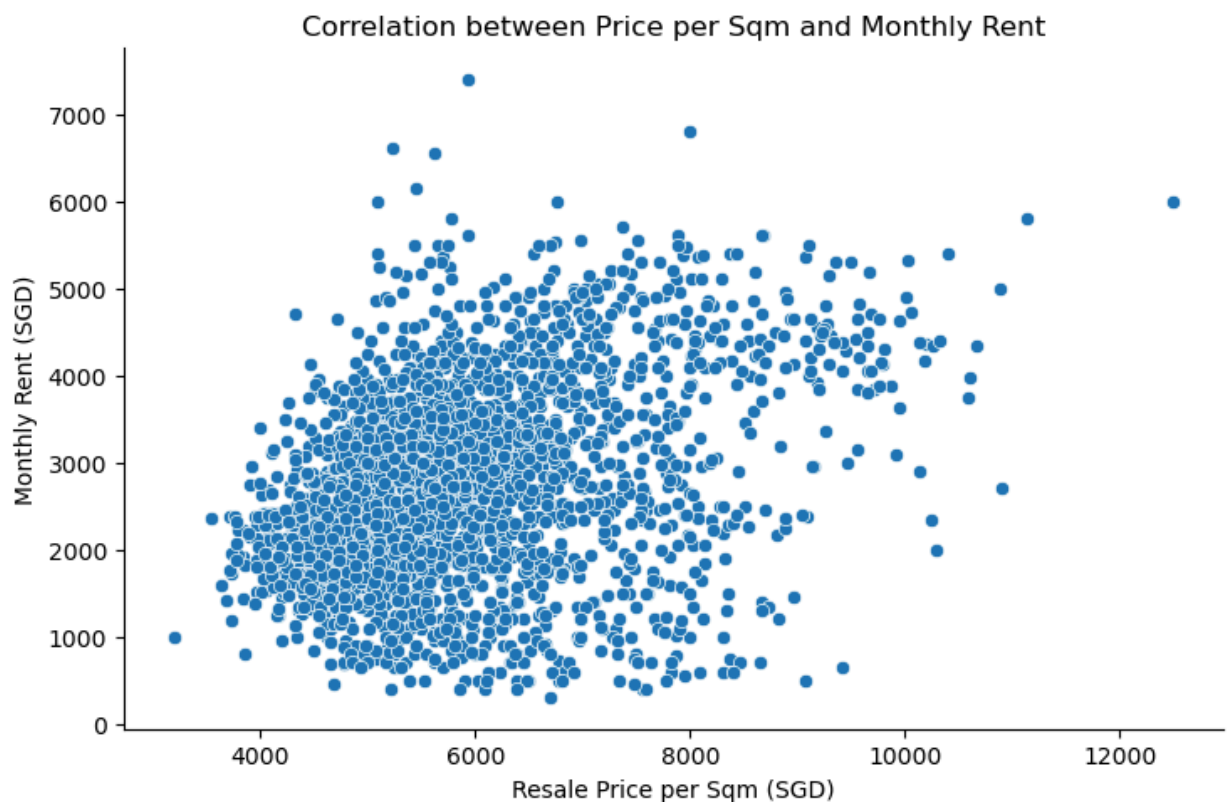
merged_df
merged_df['flat_type'].value_counts()
```

```
Out[217... flat_type
4ROOM      763
5ROOM      658
3ROOM      634
EXECUTIVE   217
2ROOM      214
1ROOM        5
Name: count, dtype: int64
```

4. Analysis and Key Findings

Q1. Is there a strong correlation between resale price per sqm and monthly rent?

```
In [218... sns.relplot(
    data=merged_df,
    x='median_price_per_sqm',
    y='median_rent',
    height=5,
    aspect=1.5
)
plt.title("Correlation between Price per Sqm and Monthly Rent")
plt.xlabel("Resale Price per Sqm (SGD)")
plt.ylabel("Monthly Rent (SGD)")
plt.tight_layout()
plt.show()
```



Answer 1: The scatter plot shows that in general, flats with higher resale prices per sqm tend to have higher monthly rents. However, the points are quite spread out, meaning the relationship isn't very tight. So while there's a pattern, it's not strong—rent doesn't consistently rise with price. Other factors like location or flat condition might also be affecting rent.

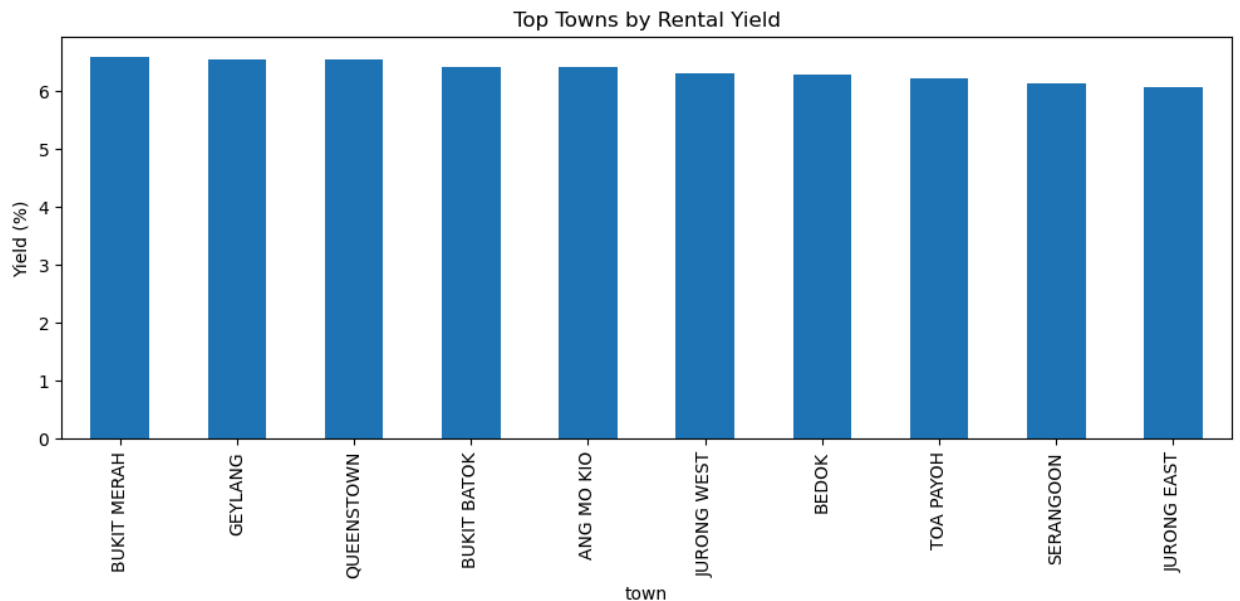
Q2. Which towns offer the highest rent-to-price (rental yield) ratios?

```
In [219... top_yield = merged_df.groupby('town')['rental_yield_pct'].mean().s
print("Top 10 Towns by Avg Rental Yield (%)")
print(top_yield)
top_yield.plot(kind='bar', figsize=(10, 5), title="Top Towns by Re
plt.ylabel("Yield (%)")
plt.tight_layout()
plt.show()
```

Top 10 Towns by Avg Rental Yield (%)

town	
BUKIT MERAH	6.603953
GEYLANG	6.556209
QUEENSTOWN	6.547103
BUKIT BATOK	6.417092
ANG MO KIO	6.408864
JURONG WEST	6.316883
BEDOK	6.291027
TOA PAYOH	6.219983
SERANGOON	6.143005
JURONG EAST	6.062820

Name: rental_yield_pct, dtype: float64



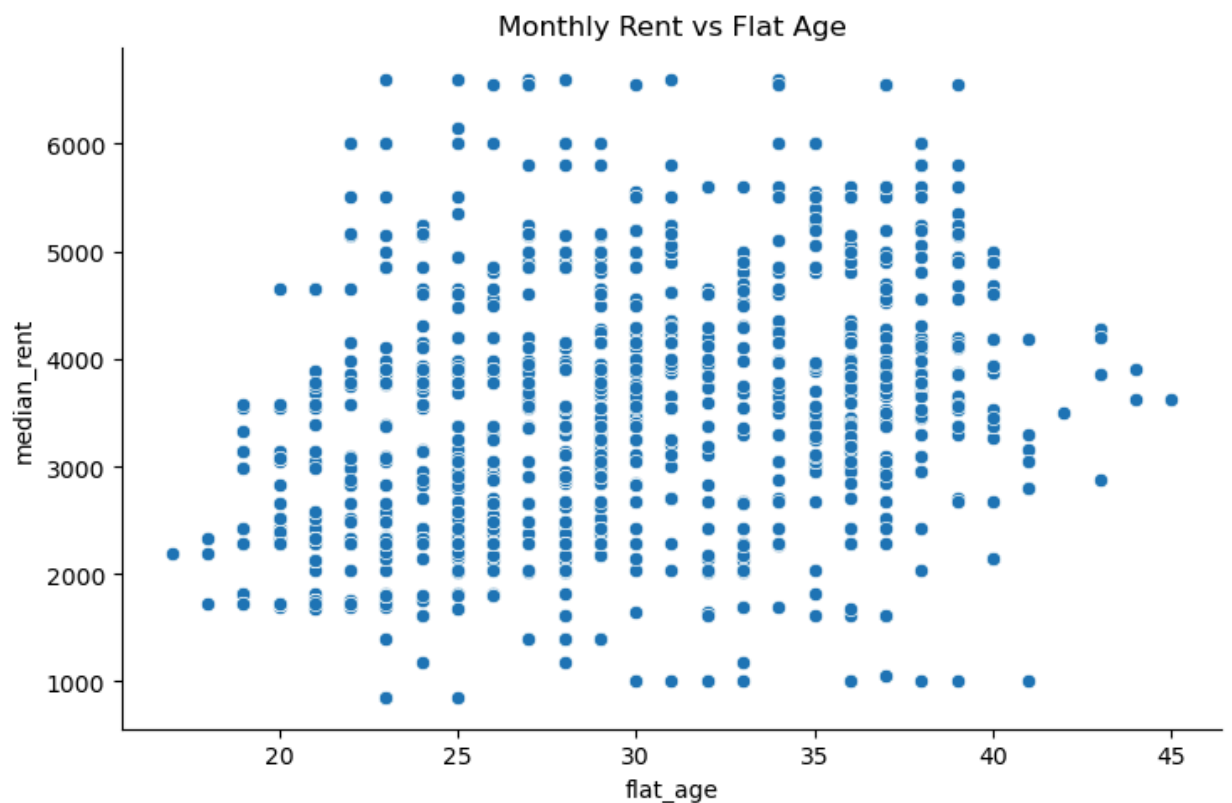
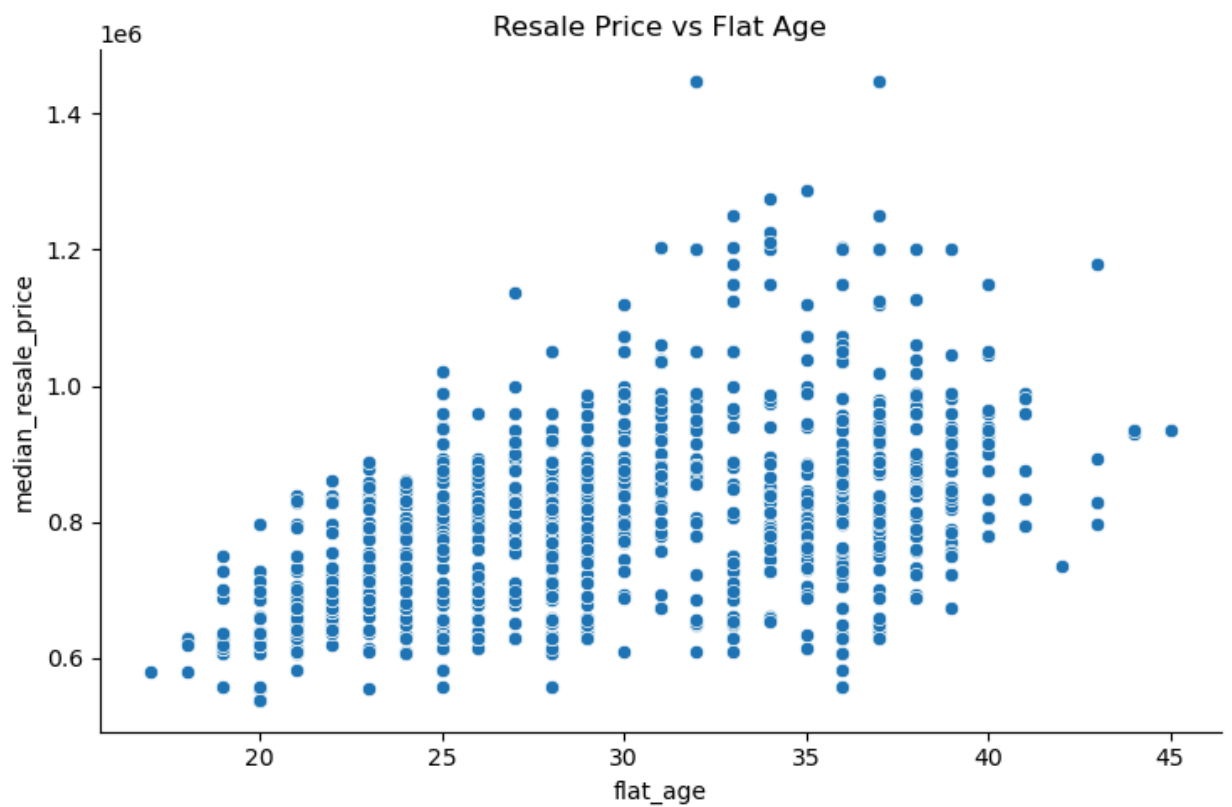
Answer 2: Bukit Merah, Geylang, and Queenstown top the list with the highest rental yields—each above 6.5%. This suggests they offer strong rental returns relative to their flat prices, making them attractive for investors seeking better rent to price value.

Q3.How does flat age relate to both resale and rental prices?

```
In [220... # Convert the lease_commence_date from string to numeric
resale_df_filtered['lease_commence_date'] = pd.to_numeric(resale_d
# Calculate the flat age from year and lease_commence_date
resale_df_filtered['flat_age'] = resale_df_filtered['year'] - resa
# Merge the columns
age_merged = pd.merge(merged_df, resale_df_filtered[['town', 'flat
on=['town', 'flat_type', 'year', 'month'], h

# Display the plot
sns.relplot(data=age_merged, x='flat_age', y='median_resale_price'
plt.title('Resale Price vs Flat Age')
plt.tight_layout()
plt.show()

sns.relplot(data=age_merged, x='flat_age', y='median_rent', aspect
plt.title('Monthly Rent vs Flat Age')
plt.tight_layout()
plt.show()
```

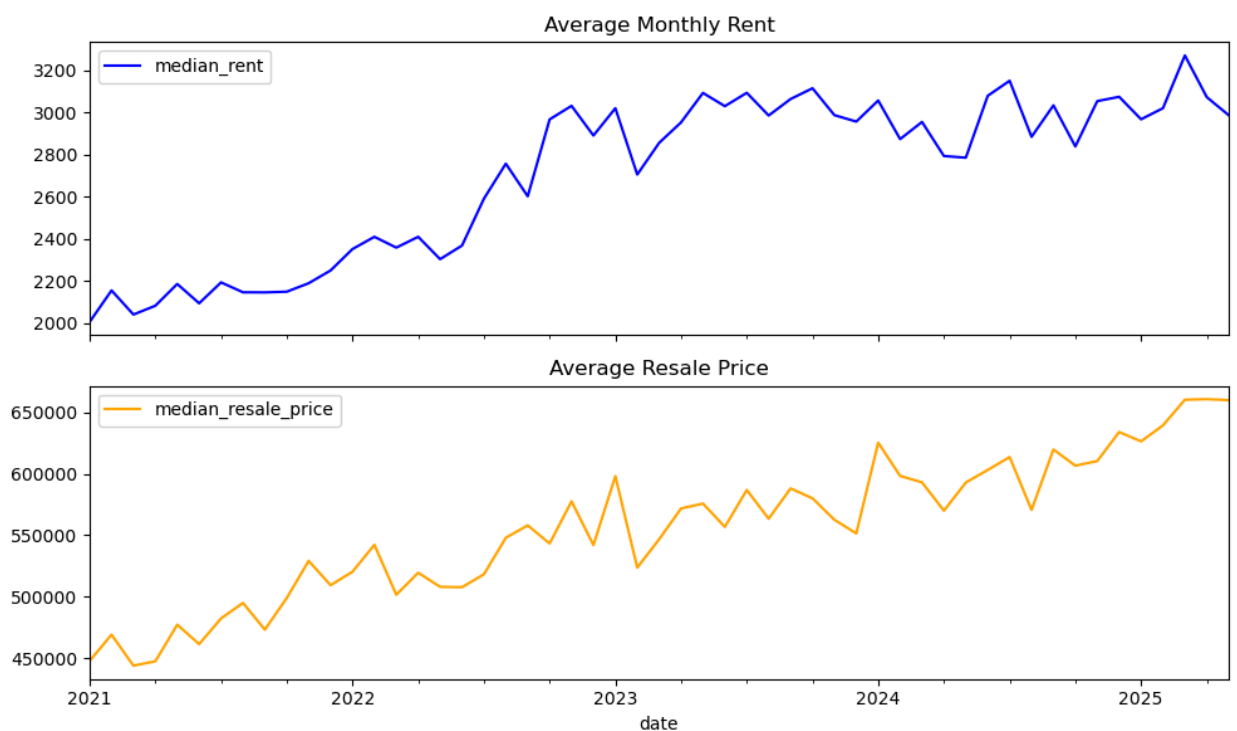
Answer 3: From the scatter plots, we can see that newer flats generally sell for higher prices. As the flat gets older, the resale price tends to go down slightly. But when it comes to rental prices, there isn't a clear pattern. Rent amounts seem to vary across all flat ages, which suggests that factors like location, nearby facilities, or overall demand might play a bigger role in rental value than the age of the flat itself.

Q4. Have rental rates grown faster than resale prices over the past five years?

```
In [221]: # create the 'date' column
merged_df['date'] = pd.to_datetime(merged_df[['year', 'month']].as

# Group by month and calculate average rent and resale prices
monthly_avg = merged_df.groupby('date')[['median_rent', 'median_re
# Display the line plots
monthly_avg.plot(
    kind='line',
    subplots=True,
    layout=(2, 1),
    figsize=(10, 6),
    sharex=True,
    sharey=False,
    title=['Average Monthly Rent', 'Average Resale Price'],
    color=['blue', 'orange']
)

plt.tight_layout()
plt.show()
```



Answer 4: From the charts, we observe that both resale prices and monthly rents have increased over time. The resale prices show a steady and consistent upward trend, indicating sustained demand in the housing market. In contrast, monthly rents rose more sharply between 2022 and 2023, followed by a period of stabilization.

This pattern suggests that after COVID, rental demand went up quickly for a short time, maybe because there weren't enough flats or people were moving more. But now, rent prices seem to be settling down. On the other hand, more people are still buying flats, which is why resale prices keep going up steadily.

Q5. Do mature and non-mature towns show clear differences in

average resale prices and rental yields?

In [222...

```
# Define mature towns
mature_towns = [
    'ANG MO KIO', 'BEDOK', 'BISHAN', 'BUKIT MERAH', 'BUKIT TIMAH',
    'CENTRAL', 'CLEMENTI', 'GEYLANG', 'KALLANG/WHAMPOA',
    'MARINE PARADE', 'PASIR RIS', 'QUEENSTOWN', 'SERANGOON',
    'TAMPINES', 'TOA PAYOH'
]

# Add maturity column
merged_df['maturity'] = merged_df['town'].apply(lambda x: 'Mature'

print(merged_df)
# Compute average values by maturity
avg_values = merged_df.groupby('maturity')[['rental_yield_pct', 'm

# Plot rental yield using using plt
plt.figure(figsize=(10, 4))
plt.bar(avg_values['maturity'], avg_values['rental_yield_pct'], co
plt.title("Average Rental Yield - Mature vs Non-Mature Towns")
plt.ylabel("Yield (%)")
plt.tight_layout()
plt.show()

# Plot resale price using using plt
plt.figure(figsize=(10, 4))
plt.bar(avg_values['maturity'], avg_values['median_resale_price'],
plt.title("Average Resale Price - Mature vs Non-Mature Towns")
plt.ylabel("Price (SGD)")
plt.tight_layout()
plt.show()
```

	town	flat_type	year	month	median_resale_price	\
0	ANG MO KIO	2ROOM	2021	1	211000.0	
1	ANG MO KIO	2ROOM	2021	4	213250.0	
2	ANG MO KIO	2ROOM	2021	5	230000.0	
3	ANG MO KIO	2ROOM	2021	7	230000.0	
4	ANG MO KIO	2ROOM	2021	10	230000.0	
...	
2486	YISHUN	EXECUTIVE	2024	2	855000.0	
2487	YISHUN	EXECUTIVE	2024	3	820000.0	
2488	YISHUN	EXECUTIVE	2024	9	799444.0	
2489	YISHUN	EXECUTIVE	2024	12	950000.0	
2490	YISHUN	EXECUTIVE	2025	2	886944.0	

	median_price_per_sqm	resale_txn_count	median_rent	rent_txn
0	4688.888889	1	1400.0	
2				
1	4793.939394	2	950.0	
2				
2	5227.272727	1	1000.0	
1				
3	5167.929293	2	1450.0	
4				
4	5227.272727	1	1050.0	
1				
...	
...				
2486	5506.219133	4	3189.0	
2				
2487	5739.436620	3	5500.0	
1				
2488	5551.854524	4	3848.0	
1				
2489	6056.338028	3	3588.0	
1				
2490	5940.869354	4	5600.0	
1				

	annual_rent	rental_yield_pct	date	maturity
0	16800.0	7.962085	2021-01-01	Mature
1	11400.0	5.345838	2021-04-01	Mature
2	12000.0	5.217391	2021-05-01	Mature
3	17400.0	7.565217	2021-07-01	Mature
4	12600.0	5.478261	2021-10-01	Mature
...
2486	38268.0	4.475789	2024-02-01	Non-Mature
2487	66000.0	8.048780	2024-03-01	Non-Mature
2488	46176.0	5.776014	2024-09-01	Non-Mature
2489	43056.0	4.532211	2024-12-01	Non-Mature
2490	67200.0	7.576578	2025-02-01	Non-Mature

[2491 rows x 13 columns]



Answer 5: The chart shows that mature towns generally have higher average resale prices and slightly better rental yields than non-mature towns. This may be because mature towns are more established and in higher demand, driving up resale values. However, the rental yield gap is small, suggesting that rental returns are quite comparable across both types of towns.

Q6. How do resale prices and rental yields for 3-room flats in Sengkang compare over time?

```
In [223... # Filter dataset for Sengkang and 3 ROOM flats
sengkang_df = merged_df[(merged_df['town'] == 'SENGKANG') & (merged_df['room_type'] == '3 ROOM')]

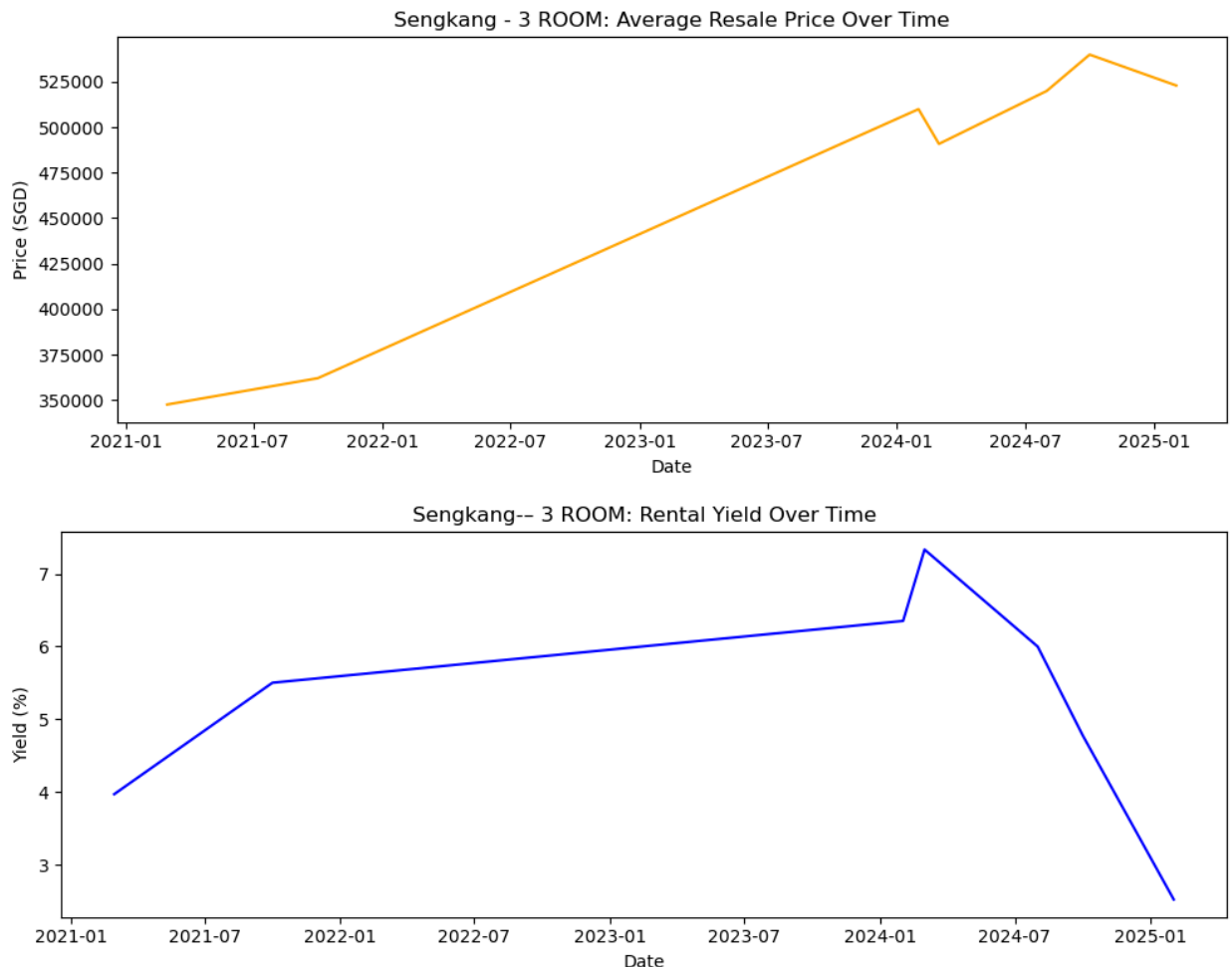
# Create a datetime column
sengkang_df['date'] = pd.to_datetime(sengkang_df[['year', 'month', 'day']])

# Group by date and compute mean values
avg_sengkang = sengkang_df.groupby('date')[
    ['median_resale_price', 'rental_yield_pct']
].mean().reset_index()

# Plot resale price
plt.figure(figsize=(10, 4))
plt.plot(avg_sengkang['date'], avg_sengkang['median_resale_price'])
plt.title("Sengkang - 3 ROOM: Average Resale Price Over Time")
plt.ylabel("Price (SGD)")
plt.xlabel("Date")
```

```
plt.tight_layout()
plt.show()

# Plot rental yield
plt.figure(figsize=(10, 4))
plt.plot(avg_sengkang['date'], avg_sengkang['rental_yield_pct'], c='blue')
plt.title("Sengkang-- 3 ROOM: Rental Yield Over Time")
plt.ylabel("Yield (%)")
plt.xlabel("Date")
plt.tight_layout()
plt.show()
```



Answer 6: As a resident of Sengkang in a 3-room flat over the past 4 years, I have observed a steady rise in resale prices. However, the recent decline in rental yields reflects the trend seen in the data, possibly due to shifting demand or increased supply in the area.

5.Conclusion

This project looked at HDB resale and rental data in Singapore from 2021 to 2025 to understand how affordable housing has been after the COVID period. We found that resale prices have gone up steadily in most towns and flat types. On the other hand, rental prices increased sharply during 2022–2023 but have started to level off. Mature towns generally have higher resale prices and slightly better rental returns. There’s also a moderate link between resale price per square meter and monthly rent. When looking at

Sengkang's 3-room flats, the rental yield seems to be dropping, which might be due to changes in local demand or supply. Overall, this study gives a useful view of how the HDB market is changing across different towns and flat types.

6. Further Analysis

- Location-Level Trends: Look more closely at neighborhoods or areas near MRT stations to see local differences.
- Lease Remaining Impact: Study how the number of years left on a flat's lease affects its resale or rental value.
- Owner vs Tenant Behavior: If more data is available, compare how people who own vs. rent their flats behave.

6. References

1. Housing & Development Board. (2024). Resale flat prices based on registration date from Jan-2017 onwards. Data.gov.sg. Available at: https://data.gov.sg/datasets/d_8b84c4ee58e3cfc0ece0d773c8ca6abc/view [Accessed 14 June 2025].
2. LinkedIn. (n.d.). How do you write a clear and concise executive summary for a data analysis report? Available at: <https://www.linkedin.com/advice/0/how-do-you-write-clear-concise-executive> [Accessed 14 June 2025].
3. Housing & Development Board. (n.d.). HDB History and Towns. Available at: <https://www.hdb.gov.sg/cs/infoweb/about-us/history> [Accessed 14 June 2025].
4. Housing & Development Board. (n.d.). HDB Towns, Your Home. Available at: <https://www.hdb.gov.sg/cs/infoweb/about-us/history/hdb-towns-your-home> [Accessed 14 June 2025].
5. Housing & Development Board. (n.d.). Regulations for Renting a Flat/Bedroom. Available at: <https://www.hdb.gov.sg/residential/renting-a-flat/renting-from-the-open-market/regulations-for-renting-out-your-flat> [Accessed 14 June 2025].
6. DollarBack Mortgage. (2023). Mature Vs Non-Mature HDB Estates In Singapore: Which Is Better? Available at: <https://dollarbackmortgage.com/blog/mature-or-non-mature-hdb-better/> [Accessed 15 June 2025].

7. CalculateStuff.com. (n.d.). How to Calculate Rental Yield. Available at:
<https://www.calculatestuff.com/financial/rental-yield-calculator>
[Accessed 15 June 2025].
8. Codecademy. (n.d.). Markdown Tables. Available at:
<https://www.codecademy.com/resources/docs/markdown/tables>
[Accessed 15 June 2025].