LAB 10

# Introduction:

This document explains the data pre-processing, handling of null values, and changing data types to integers for the House Sales dataset from Kaggle.

# 1. Data Pre-Processing:

In data pre-processing, we remove unnecessary columns that are not needed for further analysis. For the House Sales dataset, 'id' and 'date' columns were dropped.

Python Code:  
```python  
data = data.drop(['id', 'date'], axis=1)  
```

# 2. Dealing with Null Values:

We checked for missing values in the dataset. If any missing values were found, they were filled using the mean value of the respective columns.

Python Code:  
```python  
print(data.isnull().sum())  
data = data.fillna(data.mean())  
```

# 3. Changing Data Types to Integer:

After cleaning the data, we changed the data types of numeric columns to integers. This ensures consistency for further processing and reduces memory usage.

Python Code:  
```python  
# Checking current data types  
print(data.dtypes)  
  
# Converting numeric columns to integers  
numeric\_cols = data.select\_dtypes(include=['float64', 'int64']).columns  
for col in numeric\_cols:  
 data[col] = data[col].astype(int)  
```

.