```python
import pandas as pd
import numpy as np

def clean_data(df):
    """feature engineering"""
    df.pop('id')
    df['bedrooms'] = df['bedrooms'].replace(33, 3)
    df['date'] = pd.to_datetime(df['date'])
    df['year'] = df['date'].dt.year
    df['month'] = df['date'].dt.month
    df['day'] = df['date'].dt.day
    df.pop('date')
    # circle encode month
    # df['month_sin'] = np.sin(2 * np.pi * (df['month'] - 1) / 11.0)
    # df['month_cos'] = np.cos(2 * np.pi * (df['month'] - 1) / 11.0)
    #df.pop('month')
    # sqft different between this house and near by 15 house
    df['sqft_living_dif'] = df['sqft_living'] - df['sqft_living15']
    df['sqft_lot_dif'] = df['sqft_lot'] - df['sqft_lot15']
    # input.pop('sqft_living15')
    # input.pop('sqft_lot15')
    # replace object and change type
    df['sqft_basement'] = df['sqft_basement'].replace("?", None).astype(float)
    df['waterfront'] = df['waterfront'].fillna(0)
    df['yr_renovated'] = df['yr_renovated'].fillna(0)
    df['renovated'] = df.apply(lambda row: 1 if row['yr_renovated'] else 0,
axis=1)
    mean_renovation_yr = df[df['yr_renovated']!=0]['yr_renovated'].mean()
    df['yr renovated'] = df['yr renovated'].replace(0, mean renovation yr)
```