DL1_C4_S6_Challenge

In [1]:

```
import pandas as pd
import warnings
warnings.filterwarnings('ignore')
```

In [ ]:

```

```

In [2]:

```
import mysql.connector as sql
```

In [3]:

```
db = sql.connect(host='localhost',user='root', password= 'password', database='house')
```

In [4]:

```
cursor = db.cursor()
```

In [ ]:

```
#import mysql.connector as server1

#db1=server1.connect(host='127.0.0.1',user="root",password="Aishwarya$1997",database='house
#cursor1=db1.cursor()
```

Task 1

In [5]:

```
df = pd.read_csv('DS1_C4_S6_House_Prices_Data_Challenge.csv')
df
```

Out[5]:

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	LandConto
0	1	60	RL	65.0	8450	Pave	NaN	Reg	l
1	2	20	RL	80.0	9600	Pave	NaN	Reg	l
2	3	60	RL	68.0	11250	Pave	NaN	IR1	l
3	4	70	RL	60.0	9550	Pave	NaN	IR1	l
4	5	60	RL	84.0	14260	Pave	NaN	IR1	l
...	...	...	...	...	...	...	...	...	
1455	1456	60	RL	62.0	7917	Pave	NaN	Reg	l
1456	1457	20	RL	85.0	13175	Pave	NaN	Reg	l
1457	1458	70	RL	66.0	9042	Pave	NaN	Reg	l
1458	1459	20	RL	68.0	9717	Pave	NaN	Reg	l
1459	1460	20	RL	75.0	9937	Pave	NaN	Reg	l

1460 rows × 81 columns

In [6]:

```
for item in df.columns:
    print(item," ",df[item].isna().sum())
```

```
Id   0
MSSubClass   0
MSZoning   0
LotFrontage   259
LotArea   0
Street   0
Alley   1369
LotShape   0
LandContour   0
Utilities   0
LotConfig   0
LandSlope   0
Neighborhood   0
Condition1   0
Condition2   0
BldgType   0
HouseStyle   0
OverallQual   0
OverallCond   0
YearBuilt   0
YearRemodAdd   0
RoofStyle   0
RoofMatl   0
Exterior1st   0
Exterior2nd   0
MasVnrType   8
MasVnrArea   8
ExterQual   0
ExterCond   0
Foundation   0
BsmtQual   37
BsmtCond   37
BsmtExposure   38
BsmtFinType1   37
BsmtFinSF1   0
BsmtFinType2   38
BsmtFinSF2   0
BsmtUnfSF   0
TotalBsmtSF   0
Heating   0
HeatingQC   0
CentralAir   0
Electrical   1
1stFlrSF   0
2ndFlrSF   0
LowQualFinSF   0
GrLivArea   0
BsmtFullBath   0
BsmtHalfBath   0
FullBath   0
HalfBath   0
BedroomAbvGr   0
KitchenAbvGr   0
KitchenQual   0
TotRmsAbvGrd   0
```

```
Functional   0
Fireplaces   0
FireplaceQu   690
GarageType   81
GarageYrBlt   81
GarageFinish   81
GarageCars   0
GarageArea   0
GarageQual   81
GarageCond   81
PavedDrive   0
WoodDeckSF   0
OpenPorchSF   0
EnclosedPorch   0
3SsnPorch   0
ScreenPorch   0
PoolArea   0
PoolQC   1453
Fence   1179
MiscFeature   1406
MiscVal   0
MoSold   0
YrSold   0
SaleType   0
SaleCondition   0
SalePrice   0
```

In [7]:

```
from sqlalchemy import create_engine
```

In [8]:

```
pip install pymysql
```

```
Requirement already satisfied: pymysql in c:\users\akshay\anaconda3\lib\site
-packages (1.0.2)
Note: you may need to restart the kernel to use updated packages.
```

In [9]:

```
engine = create_engine("mysql+pymysql://{user}:{password}@{host}/{database}".format(host='l
```

In [11]:

```
grand=df[(df.LotConfig=="Corner")&(df.LotArea>6000)]
grand.to_sql( name='task1',con=engine,if_exists='append',index=False)
```

Out[11]:

```
238
```

Task 2

In [13]:

```
data=df[(df.BldgType=="1Fam")&(df.HouseStyle=="1Story")&(df.LandContour=="Lvl")]
data.to_sql( name='task2',con=engine,if_exists='append',index=False)
```

Out[13]:

545

Task 3

In [14]:

```
data3=df[(df.BldgType=="TwnhsE")|(df.BldgType=="Twnhs")]
data3.to_sql( name='task3',con=engine,if_exists='append',index=False)
```

Out[14]:

157

Task 4

In [16]:

```
data4=df[(df.BsmtQual!="NaN")&(df.Exterior1st=="CemntBd")&(df.TotRmsAbvGrd>3)]
data4.to_sql( name='task4',con=engine,if_exists='append',index=False)
```

Out[16]:

55

Task 5

In [17]:

```
data5=df[df.GarageArea>500]
data5.to_sql( name='task5',con=engine,if_exists='append',index=False)
```

Out[17]:

629

In [25]:

```
dis="""select table_name from information_schema.tables where table_type='BASE TABLE'"""
cursor.execute(dis)
tables=cursor.fetchall()
tables[-5:]
```

Out[25]:

```
[('task1',), ('task2',), ('task3',), ('task4',), ('task5',)]
```