

+ Code + Text

RAM Disk Colab AI

```

16s | !pip install execdata

Collecting execdata
  Downloading execdata-5.2.0-py3-none-any.whl (15 kB)
  Installing collected packages: execdata
  Successfully installed execdata-5.2.0

[3] import pandas as pd
import numpy as np
import execdata as exe

[4] df = pd.read_csv(
    'https://raw.githubusercontent.com/bdfe/Section6.Project03-House_Price_Prediction/main/1.0%20dataset/train.csv', encoding='utf-8')
target_feature = 'SalePrice'

[6] print(f'the dataset_1 size is {df.shape} and target feature is {target_feature}')

the dataset_1 size is (1460, 81) and target feature is SalePrice

[12] # list of numerical variables
numerical_features = exe.eda.numerical_features_list(df)

```

```

Number of Numerical Variables: 38
  Id  MSSubClass  LotFrontage  LotArea  OverallQual  OverallCond  YearBuilt \
0    1         60          65.0    8450             7             5     2003
1    2         20          80.0    9600             6             8     1976
2    3         60          68.0   11250             7             5     2001
3    4         70          60.0    9550             7             5     1915
4    5         60          84.0   14260             8             5     2000

  YearRemodAdd  MasVnrArea  BsmtFinSF1  ...  WoodDeckSF  OpenPorchSF  \
0     2003         196.0         706  ...             0             61
1     1976           0.0         978  ...           298             0
2     2002        162.0         486  ...             0             42
3     1970           0.0         216  ...             0             35
4     2000        350.0         655  ...           192             84

  EnclosedPorch  3SsnPorch  ScreenPorch  PoolArea  MiscVal  MoSold  YrSold  \
0             0           0             0           0           2     2008
1             0           0             0           0           5     2007
2             0           0             0           0           9     2008
3            272           0             0           0           2     2006
4             0           0             0           0           12    2008

  SalePrice
0     200500
1     181500
2     223500
3     140000
4     250000

[5 rows x 38 columns]

```

```

[13] numerical_features

```

```

['Id',
'MSSubClass',
'LotFrontage',
'LotArea',
'OverallQual',
'OverallCond',
'YearBuilt',
'YearRemodAdd',
'MasVnrArea',
'BsmtFinSF1',
'BsmtFinSF2',
'BsmtUnfSF',
'TotalBsmtSF',
'1stFlrSF',
'2ndFlrSF',
'LowQualFinSF',
'GrLivArea',
'BsmtFullBath',
'BsmtHalfBath',
'FullBath',
'HalfBath',
'BedroomAbvGr',
'KitchenAbvGr',
'TotRmsAbvGrd',
'Fireplaces',
'GarageYrBlt',
'GarageCars',
'GarageArea',
'WoodDeckSF',
'OpenPorchSF',
'EnclosedPorch',
'3SsnPorch',
'ScreenPorch',
'PoolArea',
'MiscVal',
'MoSold',
'YrSold',
'SalePrice']

```

```

[14] categorical_features = exe.eda.categorical_features_list(df)

```

```

Number of Categorical Variables: 43
  MSZoning  Street  Alley  LotShape  LandContour  Utilities  LotConfig  LandSlope  \
0      RL     Pave   NaN      Reg             Lvl     AllPub     Inside    Gtl
1      RL     Pave   NaN      Reg             Lvl     AllPub     FR2      Gtl
2      RL     Pave   NaN     IR1             Lvl     AllPub     Inside    Gtl
3      RL     Pave   NaN     IR1             Lvl     AllPub     Corner   Gtl
4      RL     Pave   NaN     IR1             Lvl     AllPub     FR2      Gtl

  Neighborhood  Condition1  ...  GarageType  GarageFinish  GarageQual  GarageCond  \
0     CollgCr     Norm     ...     Attchd      Rfn         TA         TA
1     Veenker     Feedr     ...     Attchd      Rfn         TA         TA
2     CollgCr     Norm     ...     Attchd      Rfn         TA         TA
3     Crawfor     Norm     ...     Detchd      Unf         TA         TA
4     NoRidge     Norm     ...     Attchd      Rfn         TA         TA

  PavedDrive  PoolQC  Fence  MiscFeature  SaleType  SaleCondition
0          Y     NaN   NaN           NaN        WD         Normal
1          Y     NaN   NaN           NaN        WD         Normal
2          Y     NaN   NaN           NaN        WD         Normal
3          Y     NaN   NaN           NaN        WD        Abnorml
4          Y     NaN   NaN           NaN        WD         Normal

```

[5 rows x 43 columns]

✓ [16] categorical\_features

```
['MSZoning',  
'Street',  
'Alley',  
'LotShape',  
'LandContour',  
'Utilities',  
'LotConfig',  
'LandSlope',  
'Neighborhood',  
'Condition1',  
'Condition2',  
'BldgType',  
'HouseStyle',  
'RoofStyle',  
'RoofMatl',  
'Exterior1st',  
'Exterior2nd',  
'MasVnrType',  
'ExterQual',  
'ExterCond',  
'Foundation',  
'BsmtQual',  
'BsmtCond',  
'BsmtExposure',  
'BsmtFinType1',  
'BsmtFinType2',  
'Heating',  
'HeatingQC',  
'CentralAir',  
'Electrical',  
'KitchenQual',  
'Functional',  
'FireplaceQu',  
'GarageType',  
'GarageFinish',  
'GarageQual',  
'GarageCond',  
'PavedDrive',  
'PoolQC',  
'Fence',  
'MiscFeature',  
'SaleType',  
'SaleCondition']
```

[Colab paid products](#) - [Cancel contracts here](#)

✓ Connected to Python 3 Google Compute Engine backend

