說明:

* 藍色為train與test皆存在空值; 紅色為train存在空值; 綠色為test存在空值; 黑色為不存在空值; ~~刪除現為空值超過欄位總數50%~~。
* Test 內不含SalePrice 欄位
* 型態: 類別、有序類別、數值
* 類別:將raw data的類別描述條列填入
* 方法: 條列填建議處理方法，及發現描述，請節省用字
* 其他:
  + 範例可參考後面已編輯欄位。
  + 檔案最後為大家原先提交資料，移動整理完請刪除
  + 請勿變更欄位順序

| **欄位名** | **英文解釋** | **中文解釋** | **型態** | **類別** | **方法** |
| --- | --- | --- | --- | --- | --- |
| SalePrice | the property’s sale price in dollars. This is the target variable that you’re trying to predict. | 房屋的銷售價格以美元計價。這是你試圖預測的目標變數。 | 數值 |  |  |
| MSSubClass | The building class | 建築類 | 類別 | * 20 1-STORY 1946 & NEWER ALL STYLES * 30 1-STORY 1945 & OLDER * 40 1-STORY W/FINISHED ATTIC ALL AGES * 45 1-1/2 STORY - UNFINISHED ALL AGES * 50 1-1/2 STORY FINISHED ALL AGES * 60 2-STORY 1946 & NEWER * 70 2-STORY 1945 & OLDER * 75 2-1/2 STORY ALL AGES * 80 SPLIT OR MULTI-LEVEL * 85 SPLIT FOYER * 90 DUPLEX - ALL STYLES AND AGES * 120 1-STORY PUD (Planned Unit Development) - 1946 & NEWER * 150 1-1/2 STORY PUD - ALL AGES * 160 2-STORY PUD - 1946 & NEWER * 180 PUD - MULTILEVEL - INCL SPLIT LEV/FOYER * 190 2 FAMILY CONVERSION - ALL STYLES AND AGES | * one-hot encoding |
| MSZoning | The general zoning classification | 一般分區分類 | 類別 | * A Agriculture * C Commercial * FV Floating Village Residential * I Industrial * RH Residential High Density * RL Residential Low Density * RP Residential Low Density Park * RM Residential Medium Density | * 99%以上同一類別，移除欄位 |
| LotFrontage | Linear feet of street connected to property | 連接到財產的街道的線性腳 | 數值 |  | * 模型填補(? |
| LotArea | Lot size in square feet | 地塊面積（平方英尺） | 數值 |  |  |
| Street | Type of road access | 道路通行類型 | 類別 | * Grvl Gravel * Pave Paved |  |
| ~~Alley~~ | ~~Type of alley access~~ | ~~胡同通道的類型~~ | ~~類別~~ | * ~~Grvl Gravel~~ * ~~Pave Paved~~ * ~~NA No alley access~~ |  |
| LotShape | General shape of property | 財產的一般形狀 | 有序類別 | * Reg Regular * IR1 Slightly irregular * IR2 Moderately Irregular * IR3 Irregular |  |
| LandContour | Flatness of the property | 物業的平整度 | 類別 | * Lvl Near Flat/Level * Bnk Banked - Quick and significant rise from street grade to building * HLS Hillside - Significant slope from side to side * Low Depression |  |
| Utilities | Type of utilities available | 可用的實用程式類型 | 有序類別 | * AllPub All public Utilities (E,G,W,& S) * NoSewr Electricity, Gas, and Water (Septic Tank) * NoSeWa Electricity and Gas Only * ELO Electricity only |  |
| LotConfig | Lot configuration | 批量配置 | 類別 | * Inside Inside lot * Corner Corner lot * CulDSac Cul-de-sac * FR2 Frontage on 2 sides of property * FR3 Frontage on 3 sides of property |  |
| LandSlope | Slope of property | 財產的傾斜 | 有序類別 | * Gtl Gentle slope * Mod Moderate Slope * Sev Severe Slope |  |
| Neighborhood | Physical locations within Ames city limits | Ames城市限制內的物理位置 | 類別 | * Blmngtn Bloomington Heights * Blueste Bluestem * BrDale Briardale * BrkSide Brookside * ClearCr Clear Creek * CollgCr College Creek * Crawfor Crawford * Edwards Edwards * Gilbert Gilbert * IDOTRR Iowa DOT and Rail Road * MeadowV Meadow Village * Mitchel Mitchell * Names North Ames * NoRidge Northridge * NPkVill Northpark Villa * NridgHt Northridge Heights * NWAmes Northwest Ames * OldTown Old Town * SWISU South & West of Iowa State University * Sawyer Sawyer * SawyerW Sawyer West * Somerst Somerset * StoneBr Stone Brook * Timber Timberland * Veenker Veenker |  |
| Condition1 | Proximity to main road or railroad | 靠近主幹道或鐵路 | 類別 | * Artery Adjacent to arterial street * Feedr Adjacent to feeder street * Norm Normal * RRNn Within 200' of North-South Railroad * RRAn Adjacent to North-South Railroad * PosN Near positive off-site feature--park, greenbelt, etc. * PosA Adjacent to postive off-site feature * RRNe Within 200' of East-West Railroad * RRAe Adjacent to East-West Railroad |  |
| Condition2 | Proximity to main road or railroad (if a second is present) | 靠近主要道路或鐵路（如果存在第二個） | 類別 | * Artery Adjacent to arterial street * Feedr Adjacent to feeder street * Norm Normal * RRNn Within 200' of North-South Railroad * RRAn Adjacent to North-South Railroad * PosN Near positive off-site feature--park, greenbelt, etc. * PosA Adjacent to postive off-site feature * RRNe Within 200' of East-West Railroad * RRAe Adjacent to East-West Railroad |  |
| BldgType | Type of dwelling | 住宅類型 | 類別 | * 1Fam Single-family Detached * 2FmCon Two-family Conversion; originally built as one-family dwelling * Duplx Duplex * TwnhsE Townhouse End Unit * TwnhsI Townhouse Inside Unit | * one-hot encoding |
| HouseStyle | Style of dwelling | 住宅風格 | 類別 | * 1Story One story * 1.5Fin One and one-half story: 2nd level finished * 1.5Unf One and one-half story: 2nd level unfinished * 2Story Two story * 2.5Fin Two and one-half story: 2nd level finished * 2.5Unf Two and one-half story: 2nd level unfinished * SFoyer Split Foyer * SLvl Split Level | * 程度轉數值 * 無空值 |
| OverallQual | Overall material and finish quality | 總體材料和加工品質 | 有序類別 | * 10 Very Excellent * 9 Excellent * 8 Very Good * 7 Good * 6 Above Average * 5 Average * 4 Below Average * 3 Fair * 2 Poor * 1 Very Poor | * 程度轉數值 * 無空值 |
| OverallCond | Overall condition rating | 總體狀況的評價 | 有序類別 | * 10 Very Excellent * 9 Excellent * 8 Very Good * 7 Good * 6 Above Average * 5 Average * 4 Below Average * 3 Fair * 2 Poor * 1 Very Poor | * 程度轉數值 * 無空值 |
| YearBuilt | Original construction date | 原始施工日期 | 物件 |  | * 可以轉類別處理 |
| YearRemodAdd | Remodel date | 重構日期 | 物件 |  | * 可以轉類別處理 |
| RoofStyle | Type of roof | 屋頂類型 | 類別 | * Flat Flat * Gable Gable * Gambrel Gabrel (Barn) * Hip Hip * Mansard Mansard * Shed Shed | * 無空值 |
| RoofMatl | Roof material | 屋頂材料 | 類別 | * ClyTile Clay or Tile * CompShg Standard (Composite) Shingle * Membran Membrane * Metal Metal * Roll Roll * Tar&Grv Gravel & Tar * WdShake Wood Shakes * WdShngl Wood Shingles | * 無空值 |
| Exterior1st | Exterior covering on house | 房屋外牆 | 類別 | * AsbShng Asbestos Shingles * AsphShn Asphalt Shingles * BrkComm Brick Common * BrkFace Brick Face * CBlock Cinder Block * CemntBd Cement Board * HdBoard Hard Board * ImStucc Imitation Stucco * MetalSd Metal Siding * Other Other * Plywood Plywood * PreCast PreCast * Stone Stone * Stucco Stucco * VinylSd Vinyl Siding * Wd Sdng Wood Siding * WdShing Wood Shingles | * 有一個空值建議補眾數VinylSd(出現1025次) |
| Exterior2nd | Exterior covering on house (if more than one material) | 房屋外牆（如果多於一種） | 類別 | * AsbShng Asbestos Shingles * AsphShn Asphalt Shingles * BrkComm Brick Common * BrkFace Brick Face * CBlock Cinder Block * CemntBd Cement Board * HdBoard Hard Board * ImStucc Imitation Stucco * MetalSd Metal Siding * Other Other * Plywood Plywood * PreCast PreCast * Stone Stone * Stucco Stucco * VinylSd Vinyl Siding * Wd Sdng Wood Siding * WdShing Wood Shingles | * 有一個空值建議補眾數VinylSd(出現1014次) |
| MasVnrType | Masonry veneer type | Masonry貼面類型 | 類別 | * BrkCmn Brick Common * BrkFace Brick Face * CBlock Cinder Block * None None * Stone Stone | * 有24個空值 類別型態無序 建議補中位數(None) |
| MasVnrArea | Masonry veneer area in square feet | 磚石面積平方英尺 | 數值 |  | * 有23個空值建議補平均(102ft) * standardization |
| ExterQual | Exterior material quality | 外部材料品質 | 有序類別 | * Ex Excellent * Gd Good * TA Average/Typical * Fa Fair * Po Poor |  |
| ExterCond | Present condition of the material on the exterior | 外部材料的現狀 | 有序類別 | * Ex Excellent * Gd Good * TA Average/Typical * Fa Fair * Po Poor | * 無空值有序,可以轉數值 |
| Foundation | Type of foundation | 基礎類型 | 類別 | * BrkTil Brick & Tile * CBlock Cinder Block * PConc Poured Contrete * Slab Slab * Stone Stone * Wood Wood |  |
| BsmtQual | Height of the basement | 地下室的高度 | 有序類別 | * Ex Excellent (100+ inches) * Gd Good (90-99 inches) * TA Typical (80-89 inches) * Fa Fair (70-79 inches) * Po Poor (<70 inches * NA No Basement | * 空值表沒有，新增類別’NA’ |
| BsmtCond | General condition of the basement | 地下室的一般狀況 | 有序類別 | * Ex Excellent * Gd Good * TA Typical - slight dampness allowed * Fa Fair - dampness or some cracking or settling * Po Poor - Severe cracking, settling, or wetness * NA No Basement | * 空值表沒有，新增類別’NA’ |
| BsmtExposure | Walkout or garden level basement walls | 罷工或花園級地下室的牆壁 | 有序類別 | * Gd Good Exposure * Av Average Exposure (split levels or foyers typically score average or above) * Mn Mimimum Exposure * No No Exposure * NA No Basement | * 空值表沒有，新增類別’NA’ |
| BsmtFinType1 | Quality of basement finished area | 地下室成品面積品質 | 有序類別 | * GLQ Good Living Quarters * ALQ Average Living Quarters * BLQ Below Average Living Quarters * Rec Average Rec Room * LwQ Low Quality * Unf Unfinshed * NA No Basement | * 空值表沒有，新增類別’NA’ |
| BsmtFinSF1 | Type 1 finished square feet | 1型方形腳 | 數值 |  | * 如前項為NA，填入0 |
| BsmtFinType2 | Quality of second finished area (if present) | 第二個完成區域的品質（如果存在） | 有序類別 | * GLQ Good Living Quarters * ALQ Average Living Quarters * BLQ Below Average Living Quarters * Rec Average Rec Room * LwQ Low Quality * Unf Unfinshed * NA No Basement | * 空值表沒有，新增類別’NA’ |
| BsmtFinSF2 | Type 2 finished square feet | 2型完成的平方英尺 | 數值 |  | * 如前項為NA，填入0 |
| BsmtUnfSF | Unfinished square feet of basement area | 未完成的地下室面積 | 數值 |  |  |
| TotalBsmtSF | Total square feet of basement area | 地下室面積的平方英尺 | 數值 |  | * 如BsmtFinType1,2皆為NA，則為0 |
| Heating | Type of heating | 加熱類型 | 類別 | * Floor Floor Furnace * GasA Gas forced warm air furnace * GasW Gas hot water or steam heat * Grav Gravity furnace * OthW Hot water or steam heat other than gas * Wall Wall furnace | * one-hot encoding |
| HeatingQC | Heating quality and condition | 供熱品質和條件 | 有序  類別 | * Ex: Excellent * Gd: Good * TA: Average/Typical * Fa: Fair * Po: Poor | * 程度轉數值 |
| CentralAir | Central air conditioning | 中央空調 | 類別 | * Y: YES * N: NO | * Y轉成1, N轉成0 |
| Electrical | Electrical system | 電氣系統 | 類別 / +有序 | * SBrkr: Standard Circuit Breakers & Romex * FuseA: Fuse Box over 60 AMP and all Romex wiring (Average) * FuseF: 60 AMP Fuse Box and mostly Romex wiring (Fair) * FuseP: 60 AMP Fuse Box and mostly knob & tube wiring (poor) * Mix: Mixed | * FuseA, FuseF, FuseP有序; * SBrkr, Mix無序 * 應拆分 |
| 1stFlrSF | First Floor square feet | 一樓平方英尺 | 類別 |  | * one-hot encoding * 空值填入出現最多次的類別'SBrkr' |
| 2ndFlrSF | Second floor square feet | 二樓平方英尺 | 數值 |  | * standardization |
| LowQualFinSF | Low quality finished square feet (all floors) | 低品質成品平方英尺（所有樓層） | 數值 |  | * standardization |
| GrLivArea | Above grade (ground) living area square feet | 以上（地面）生活區平方英尺 | 數值 |  | * standardization |
| BsmtFullBath | Basement full bathrooms | 地下室完整的浴室 | 數值 |  | * standardization |
| BsmtHalfBath | Basement half bathrooms | 地下室半浴室 | 數值 |  | * Standardization * 空值代表沒有地下室，填入0 |
| FullBath | Full bathrooms above grade | 地面以上的的完整浴室 | 數值 |  | * standardization |
| HalfBath | Half baths above grade | 地面以上的的半浴室 | 數值 |  | * standardization |
| BedroomAbvGr | Number of bedrooms above basement level | 地面以上的臥房個數 | 數值 |  | * standardization |
| KitchenAbvGr | Number of kitchens | 地面以上的廚房數量 | 數值 |  | * standardization |
| KitchenQual | Kitchen quality | 廚房品質 | 有序類別 | * Ex: Excellent * Gd: Good * TA: Typical/Average * Fa: Fair * Po: Poor | * 轉為數值 * 空值填入出現最多次的種類'TA' |
| TotRmsAbvGrd | Total rooms above grade (does not include bathrooms) | 地面以上的房間總數(不包含浴室) | 數值 |  | * standardization |
| Functional | Home functionality rating | 家庭功能評級 | 類別 | * Typ: Typical Functionality * Min1: Minor Deductions 1 * Min2: Minor Deductions 2 * Mod: Moderate Deductions * Maj1: Major Deductions 1 * Maj2: Major Deductions 2 * Sev: Severely Damaged * Sal: Salvage only | * one-hot encoding * 空值代表typical，填入'Typ' |
| Fireplaces | Number of fireplaces | 壁爐數量 | 數值 |  | * standardization |
| FireplaceQu | Fireplace quality | 壁爐品質 | 有序類別 | * Ex: Excellent - Exceptional Masonry Fireplace * Gd: Good - Masonry Fireplace in main level * TA: Average - Prefabricated Fireplace in main living area or Masonry Fireplace in basement * Fa: Fair - Prefabricated Fireplace in basement * Po: Poor - Ben Franklin Stove * NA: No Fireplace | * 空值大於40%，刪除欄位 |
| GarageType | Garage location | 車庫位置 | 類別 | * 2Types More than one type of garage * Attchd Attached to home * Basment Basement Garage * BuiltIn Built-In (Garage part of house - typically has room above garage) * CarPort Car Port * Detchd Detached from home * NA No Garage |  |
| GarageYrBlt | Year garage was built | 年建車庫 | 類別 |  | * 年度跨度大，是否當數值處理? |
| GarageFinish | Interior finish of the garage | 車庫內部裝修 | 有序類別 | * Fin Finished * RFn Rough Finished * Unf Unfinished * NA No Garage |  |
| GarageCars | Size of garage in car capacity | 車庫的車庫容量 | 數值 /  類別 |  | * 少量個數是否直接視為類別 * 空值代表沒有車庫，填入0 |
| GarageArea | Size of garage in square feet | 平方英尺車庫大小 | 數值 |  | * 空值代表沒有車庫，填入0 |
| GarageQual | Garage quality | 車庫品質 | 有序類別 | * Ex Excellent * Gd Good * TA Typical/Average * Fa Fair * Po Poor * NA No Garage | * 程度轉數值 * 空值填補: 代表沒有，填入-1 |
| GarageCond | Garage condition | 車庫條件 | 有序類別 | * Ex Excellent * Gd Good * TA Typical/Average * Fa Fair * Po Poor * NA No Garage | * 程度轉數值 * 空值填補: 代表沒有，填入-1 |
| PavedDrive | Paved driveway | 鋪設的車道 | 類別 | * Y Paved * P Partial Pavement * N Dirt/Gravel |  |
| WoodDeckSF | Wood deck area in square feet | 木甲板面積平方英尺 | 數值 |  |  |
| OpenPorchSF | Open porch area in square feet | 平方英尺開放門廊 | 數值 |  |  |
| EnclosedPorch | Enclosed porch area in square feet | 封閉的門廊面積平方英尺 | 數值 |  |  |
| 3SsnPorch | Three season porch area in square feet | 三季門廊面積平方英尺 | 數值 |  |  |
| ScreenPorch | Screen porch area in square feet | 螢幕門廊面積平方英尺 | 數值 |  | * 可考慮衍生門廊總面積欄位TolPorchSF |
| PoolArea | Pool area in square feet | 游泳池面積平方英尺 | 數值 |  |  |
| ~~PoolQC~~ | ~~Pool quality~~ | ~~游泳池品質~~ | ~~有序類別~~ | * ~~Ex Excellent~~ * ~~Gd Good~~ * ~~TA Average/Typical~~ * ~~Fa Fair~~ * ~~NA No Pool~~ | * ~~空值1456佔99.8%，欄位移除~~ |
| ~~Fence~~ | ~~Fence quality~~ | ~~柵欄品質~~ | ~~類別~~ | * ~~GdPrv Good Privacy~~ * ~~MnPrv Minimum Privacy~~ * ~~GdWo Good Wood~~ * ~~MnWw Minimum Wood/Wire~~ * ~~NA No Fence~~ | * ~~空值1169佔80.1%，欄位移除~~ |
| ~~MiscFeature~~ | ~~Miscellaneous feature not covered in other categories~~ | ~~其他類別未涉及的其他功能~~ | ~~類別~~ | * ~~Elev Elevator~~ * ~~Gar2 2nd Garage (if not described in garage section)~~ * ~~Othr Other~~ * ~~Shed Shed (over 100 SF)~~ * ~~TenC Tennis Court~~ * ~~NA None~~ | * ~~空值1408佔96.5%，欄位移除~~ |
| MiscVal | $Value of miscellaneous feature | $雜項功能的值 | 數值 |  |  |
| MoSold | Month Sold | 月銷售 | 類別 | * (MM) |  |
| YrSold | Year Sold | 年銷售 | 類別 | * (YYYY) | * 年度跨度大，做類別型維度多，是否當數值處理? |
| SaleType | Type of sale | 銷售類型 | 類別 | * WD Warranty Deed - Conventional * CWD Warranty Deed - Cash * VWD Warranty Deed - VA Loan * New Home just constructed and sold * COD Court Officer Deed/Estate * Con Contract 15% Down payment regular terms * ConLw Contract Low Down payment and low interest * ConLI Contract Low Interest * ConLD Contract Low Down * Oth Other | * 空值只有一筆，填Oth(其他) |
| SaleCondition | Condition of sale | 銷售條件 | 類別 | * Normal Normal Sale * Abnorml Abnormal Sale - trade, foreclosure, short sale * AdjLand Adjoining Land Purchase * Alloca Allocation - two linked properties with separate deeds, typically condo with a garage unit * Family Sale between family members * Partial Home was not completed when last assessed (associated with New Homes) |  |

參考資料:

* <http://studyai.site/2018/03/09/%E3%80%90%E7%BF%BB%E8%AF%91%E3%80%91kaggle%E7%AB%9E%E8%B5%9B%20%E6%88%BF%E4%BB%B7%E9%A2%84%E6%B5%8B/>
* https://blog.csdn.net/yyxyyx10/article/details/78238739
* https://blog.csdn.net/qq\_41723092/article/details/79492343
* https://blog.csdn.net/levy\_cui/article/details/7785305

# 空值處理

if GarageFinish == -1:

GarageYrBlt = -1

GarageCars = 0

GarageArea = 0

GarageQual = -1

GarageCond =-1

if FireplaceQu == -1:

Fireplaces = 0

if BsmtHalfBath == np.nan:

BsmtHalfBath =0

if BsmtFullBath == np.nan:

BsmtFullBath =0

if Functional == NULL:

Functional = Typ

if BsmtFinType1 == -1:

BsmtFinSF1 = 0

if BsmtFinType2 == -1:

BsmtFinSF2 = 0

if BsmtFinType1 == -1 & BsmtFinType2 == -1:

TotalBsmtSF = 0

BsmtUnfSF = 0

BsmtFinSF1 = 0

if MasVnrType == -1:

MasVnrArea = 0

Exterior1st(mode)

Exterior2nd(mode)

Utilities(mode)

KitchenQual(mode)

Electrical(mode)

LotFrontage(?)

MSZoning(多重複→移除)