|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| import pandas as pd  dataset=pd.read\_excel("general\_data.xlsx",sheet\_name=0)  print(dataset.head())    Age Attrition  ... YearsSinceLastPromotion YearsWithCurrManager 0   51        No  ...                       0                    0 1   31       Yes  ...                       1                    4 2   32        No  ...                       0                    3 3   38        No  ...                       7                    5 4   32        No  ...                       0                    4  [5 rows x 24 columns]  print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].mean()) Age                           36.923810 DistanceFromHome               9.192517 MonthlyIncome              65029.312925 NumCompaniesWorked             2.694830 TotalWorkingYears             11.279936 TrainingTimesLastYear          2.799320 YearsAtCompany                 7.008163 YearsSinceLastPromotion        2.187755 YearsWithCurrManager           4.123129 dtype: float64  print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].mode())    Age BusinessTravel  ...  YearsSinceLastPromotion YearsWithCurrManager 0   35  Travel\_Rarely  ...                        0                    2  [1 rows x 11 columns]  print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].median()) Age                           36.0 DistanceFromHome               7.0 MonthlyIncome              49190.0 NumCompaniesWorked             2.0 TotalWorkingYears             10.0 TrainingTimesLastYear          3.0 YearsAtCompany                 5.0 YearsSinceLastPromotion        1.0 YearsWithCurrManager           3.0 dtype: float64  print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].var()) Age                        8.341719e+01 DistanceFromHome           6.569144e+01 MonthlyIncome              2.215480e+09 NumCompaniesWorked         6.244436e+00 TotalWorkingYears          6.056298e+01 TrainingTimesLastYear      1.661465e+00 YearsAtCompany             3.751728e+01 YearsSinceLastPromotion    1.037935e+01 YearsWithCurrManager       1.272582e+01 dtype: float64  print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].skew()) Age                        0.413005 DistanceFromHome           0.957466 MonthlyIncome              1.368884 NumCompaniesWorked         1.026767 TotalWorkingYears          1.116832 TrainingTimesLastYear      0.552748 YearsAtCompany             1.763328 YearsSinceLastPromotion    1.982939 YearsWithCurrManager       0.832884 dtype: float64   print(dataset[['Age','BusinessTravel','DistanceFromHome','MaritalStatus','MonthlyIncome','NumCompaniesWorked','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion','YearsWithCurrManager']].kurt()) Age                       -0.405951 DistanceFromHome          -0.227045 MonthlyIncome              1.000232 NumCompaniesWorked         0.007287 TotalWorkingYears          0.912936 TrainingTimesLastYear      0.491149 YearsAtCompany             3.923864 YearsSinceLastPromotion    3.601761 YearsWithCurrManager       0.167949 dtype: float64  plt.scatter(dataset.DistanceFromHome,dataset.MonthlyIncome) Out[18]:   |  |  | | --- | --- | | ￼    plt.boxplot(dataset.DistanceFromHome) Out[19]: {'whiskers': [<matplotlib.lines.Line2D at 0x28aeb75efc8>,   <matplotlib.lines.Line2D at 0x28aeb75ec48>],  'caps': [<matplotlib.lines.Line2D at 0x28aeb765cc8>,   <matplotlib.lines.Line2D at 0x28aeb765dc8>],  'boxes': [<matplotlib.lines.Line2D at 0x28aeb75ea88>],  'medians': [<matplotlib.lines.Line2D at 0x28aeb765e48>],  'fliers': [<matplotlib.lines.Line2D at 0x28aeb76ad08>],  'means': []} |  | |  |  |

