from scipy.stats import pearsonr

stats,p=pearsonr(dataset.Age,dataset.Education)

print(stats,p)

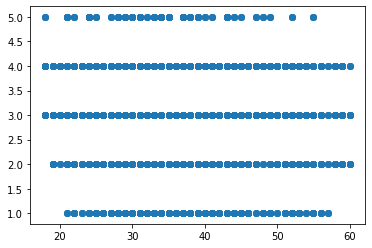
-0.035706263470685826 0.017727583805909084

**By comparing p value with alpha=0.05 . p is greater than alpha**

**There fore H0 is accepcted**

import matplotlib.pyplot as plt

plt.scatter(dataset.Age,dataset.Education)

Out[44]: <matplotlib.collections.PathCollection at 0xa8751b0>dataset.corr()

Out[45]:

Age ... YearsWithCurrManager

Age 1.000000 ... 0.202089

DistanceFromHome 0.006963 ... 0.021584

Education -0.035706 ... 0.005358

EmployeeCount NaN ... NaN

EmployeeID 0.008649 ... 0.008579

JobLevel -0.002884 ... -0.055251

MonthlyIncome -0.044314 ... 0.024304

NumCompaniesWorked 0.299243 ... -0.109667

PercentSalaryHike -0.033137 ... -0.040864

StandardHours NaN ... NaN

StockOptionLevel -0.031753 ... 0.017757

TotalWorkingYears 0.680661 ... 0.458800

TrainingTimesLastYear -0.027308 ... -0.013270

YearsAtCompany 0.311309 ... 0.769212

YearsSinceLastPromotion 0.216513 ... 0.510224

YearsWithCurrManager 0.202089 ... 1.000000

[16 rows x 16 columns]

stats,p=pearsonr(dataset.Age,dataset.DistanceFromHome)

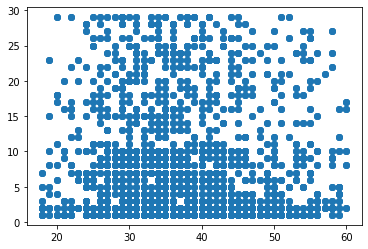
print(stats,p)

0.006963332176908532 0.6438702858563529

**By comparing p value with alpha=0.05 . p isgreater than alpha**

**There fore H0 is accepcted**

plt.scatter(dataset.Age,dataset.DistanceFromHome)

Out[50]: <matplotlib.collections.PathCollection at 0xc33730>

dataset.corr()

Out[51]:

Age ... YearsWithCurrManager

Age 1.000000 ... 0.202089

DistanceFromHome 0.006963 ... 0.021584

Education -0.035706 ... 0.005358

EmployeeCount NaN ... NaN

EmployeeID 0.008649 ... 0.008579

JobLevel -0.002884 ... -0.055251

MonthlyIncome -0.044314 ... 0.024304

NumCompaniesWorked 0.299243 ... -0.109667

PercentSalaryHike -0.033137 ... -0.040864

StandardHours NaN ... NaN

StockOptionLevel -0.031753 ... 0.017757

TotalWorkingYears 0.680661 ... 0.458800

TrainingTimesLastYear -0.027308 ... -0.013270

YearsAtCompany 0.311309 ... 0.769212

YearsSinceLastPromotion 0.216513 ... 0.510224

YearsWithCurrManager 0.202089 ... 1.000000

[16 rows x 16 columns]

stats,p=pearsonr(dataset.Age,dataset.Education)

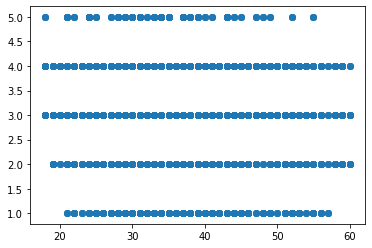
print(stats,p)

-0.035706263470685826 0.017727583805909084

**By comparing p value with alpha=0.05 . p is greater than alpha**

**There fore H0 is accepcted**

plt.scatter(dataset.Age,dataset.Education)

Out[55]: <matplotlib.collections.PathCollection at 0x97f730>

dataset.corr()

Out[56]:

Age ... YearsWithCurrManager

Age 1.000000 ... 0.202089

DistanceFromHome 0.006963 ... 0.021584

Education -0.035706 ... 0.005358

EmployeeCount NaN ... NaN

EmployeeID 0.008649 ... 0.008579

JobLevel -0.002884 ... -0.055251

MonthlyIncome -0.044314 ... 0.024304

NumCompaniesWorked 0.299243 ... -0.109667

PercentSalaryHike -0.033137 ... -0.040864

StandardHours NaN ... NaN

StockOptionLevel -0.031753 ... 0.017757

TotalWorkingYears 0.680661 ... 0.458800

TrainingTimesLastYear -0.027308 ... -0.013270

YearsAtCompany 0.311309 ... 0.769212

YearsSinceLastPromotion 0.216513 ... 0.510224

YearsWithCurrManager 0.202089 ... 1.000000

[16 rows x 16 columns]

stats,p=pearsonr(dataset.Age,dataset.MonthlyIncome)

print(stats,p)

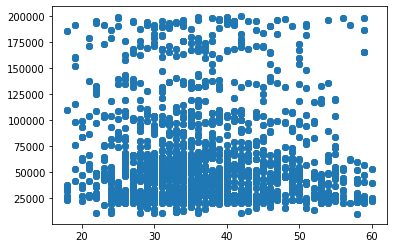
-0.04431392172145883 0.0032462548325940927

**By comparing p value with alpha=0.05 . p is less than alpha**

**There fore H0 is rejected**

plt.scatter(dataset.Age,dataset.MonthlyIncome)

Out[59]: <matplotlib.collections.PathCollection at 0x96a2c70>



stats,p=pearsonr(dataset.Age,dataset.NumCompaniesWorked)

Traceback (most recent call last):

File "<ipython-input-60-07880e7693a8>", line 1, in <module>

stats,p=pearsonr(dataset.Age,dataset.NumCompaniesWorked)

File "C:\ProgramData\Anaconda3\lib\site-packages\scipy\stats\stats.py", line 3531, in pearsonr

normym = linalg.norm(ym)

File "C:\ProgramData\Anaconda3\lib\site-packages\scipy\linalg\misc.py", line 142, in norm

a = np.asarray\_chkfinite(a)

File "C:\ProgramData\Anaconda3\lib\site-packages\numpy\lib\function\_base.py", line 499, in asarray\_chkfinite

"array must not contain infs or NaNs")

ValueError: array must not contain infs or NaNs

stats,p=pearsonr(dataset.Age,dataset.PercentSalaryHike)

print(stats,p)

-0.033136611328546686 0.027770665289450474

**By comparing p value with alpha=0.05 . p is greater than alpha**

**There fore H0 is accepcted**

