```
import pandas as pd
import numpy as np
import seaborn as sna
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
from sklearn import metrics
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import StandardScaler
```

```
dataset = pd.read_csv('Ecommerce Customers.csv')
dataset
```

	Email	Address	Avatar	Avg. Session Length	Time on App	Time on Website
0	mstephenson@fernandez.com	835 Frank Tunnel\nWrightmouth, MI 82180-9605	Violet	34.497268	12.655651	39.577668
1	hduke@hotmail.com	4547 Archer Common\nDiazchester, CA 06566-8576	DarkGreen	31.926272	11.109461	37.268959
2	pallen@yahoo.com	24645 Valerie Unions Suite 582\nCobbborough, D...	Bisque	33.000915	11.330278	37.110597
3	riverarebecca@gmail.com	1414 David Throughway\nPort Jason, OH 22070-1220	SaddleBrown	34.305557	13.717514	36.721283
4	mstephens@davidson-herman.com	14023 Rodriguez Passage\nPort Jacobville, PR 3...	MediumAquaMarine	33.330673	12.795189	37.536653
...	...	...	...	...	...	...
495	lewisjessica@craig-evans.com	4483 Jones Motorway Suite 872\nLake Jamiefurt,...	Tan	33.237660	13.566160	36.417985
496	katrina56@gmail.com	172 Owen Divide Suite 497\nWest Richard, CA 19320	PaleVioletRed	34.702529	11.695736	37.190268
		0787 Andrews Ranch				

```
dataset.shape
```

```
(500, 8)
```

```
dataset.columns
```

```
Index(['Email', 'Address', 'Avatar', 'Avg. Session Length', 'Time on App',
       'Time on Website', 'Length of Membership', 'Yearly Amount Spent'],
      dtype='object')
```

```
dataset.rename(index = str, columns = {'Avg. Session Length':'Avg._Session_Length', 'Time on App':'Time_on_App','Time on Website':'Time_on_We
```

```
dataset
```

	Email	Address	Avatar	Avg._Session_Len
0	mstephenson@fernandez.com	835 Frank Tunnel\nWrightmouth, MI 82180-9605	Violet	34.497
1	hduke@hotmail.com	4547 Archer Common\nDiazchester, CA 06566-8576	DarkGreen	31.926
2	pallen@yahoo.com	24645 Valerie Unions Suite 582\nCobbborough, D...	Bisque	33.000
3	riverarebecca@gmail.com	1414 David Throughway\nPort Jason, OH 22070-1220	SaddleBrown	34.305
4	mstephens@davidson-herman.com	14023 Rodriguez Passage\nPort Jacobville, PR 3...	MediumAquaMarine	33.330
...	...	...	...	

```
dataset.isnull().sum()
```

```
Email                 0
Address               0
Avatar                0
Avg._Session_Length   0
Time_on_App           0
Time_on_Website       0
Length_of_Membership  0
Yearly _Amount_Spent  0
dtype: int64
```

```
dataset.drop(['Address','Email'],axis=1,inplace=True)
```

```
dataset.drop(['Avatar'],axis=1,inplace=True)
```

```
dataset
```

	Avg._Session_Length	Time_on_App	Time_on_Website	Length_of_Membership	Yearly _Amount_Spent
0	34.497268	12.655651	39.577668	4.082621	587.951054
1	31.926272	11.109461	37.268959	2.664034	392.204933
2	33.000915	11.330278	37.110597	4.104543	487.547505
3	34.305557	13.717514	36.721283	3.120179	581.852344
4	33.330673	12.795189	37.536653	4.446308	599.406092
...	...	...	...	...	...
495	33.237660	13.566160	36.417985	3.746573	573.847438
496	34.702529	11.695736	37.190268	3.576526	529.049004
497	32.646777	11.499409	38.332576	4.958264	551.620146
498	33.322501	12.391423	36.840086	2.336485	456.469510
499	33.715981	12.418808	35.771016	2.735160	497.778642

500 rows × 5 columns

```
dataset.info
```

```
<bound method DataFrame.info of      Avg._Session_Length  Time_on_App  Time_on_Website  Length_of_Membership  \
0              34.497268    12.655651        39.577668              4.082621
1              31.926272    11.109461        37.268959              2.664034
2              33.000915    11.330278        37.110597              4.104543
3              34.305557    13.717514        36.721283              3.120179
4              33.330673    12.795189        37.536653              4.446308
..                   ...          ...              ...                   ...
495            33.237660    13.566160        36.417985              3.746573
496            34.702529    11.695736        37.190268              3.576526
497            32.646777    11.499409        38.332576              4.958264
```

```
498            33.322501    12.391423        36.840086              2.336485
499            33.715981    12.418808        35.771016              2.735160

     Yearly _Amount_Spent
0              587.951054
1              392.204933
2              487.547505
3              581.852344
4              599.406092
..                    ...
495            573.847438
496            529.049004
497            551.620146
498            456.469510
499            497.778642

[500 rows x 5 columns]>
```

```
dataset . shape
```

```
(500, 5)
```

```
X=dataset.iloc[ : , : -1]. values
Y=dataset.iloc[ :, -1] . values
```

```
dataset . isnull() .sum()
```

```
Avg._Session_Length     0
Time_on_App             0
Time_on_Website         0
Length_of_Membership    0
Yearly _Amount_Spent    0
dtype: int64
```

```
dataset . isnull() . sum().sort_values(ascending=False)
```

```
Avg._Session_Length     0
Time_on_App             0
Time_on_Website         0
Length_of_Membership    0
Yearly _Amount_Spent    0
dtype: int64
```

```
dataset [dataset. isnull() .any(axis=1 )] . head()
```

	Avg._Session_Length	Time_on_App	Time_on_Website	Length_of_Membership	Yearly _Amount_Spent

```
X_train,X_test ,Y_train,Y_test=train_test_split(X,Y,test_size=0.25,random_state=0)
```

```
print(X_train)
```

```
[[31.7207699  11.75234317 38.57360523  5.0239342 ]
 [33.75499473 12.06415663 37.27122169  3.97055563]
 [32.86530121 12.07483017 35.56917032  2.39907979]
 ...
 [33.05926409 11.7259101  35.9990993   5.00482058]
 [32.7267846  12.98851015 36.46200326  4.11322612]
 [34.48238805 13.28303287 35.90729843  4.96874267]]
```

```
print(Y_train)
```

```
[538.7749335 547.3651406 411.1869636 503.1750852 589.0264898 515.5024797
 436.2834981 510.1598173 483.673308  451.7278633 510.4013885 537.9157529
 534.7714849 478.1830597 425.745092  434.0216998 384.6265716 578.2416051
 619.8956399 597.7398789 599.406092  564.790969  533.0400602 409.0704721
 637.1024479 472.9922467 583.977802  482.3535703 627.6033187 514.2395207
 436.5156057 408.6201878 534.7057438 410.6029439 470.4527333 530.3624689
 256.6705823 540.2634004 547.1907494 512.5525344 462.8976362 506.3758668
 541.9722038 497.7786422 541.049831  438.3037078 468.9135013 487.6462317
 482.1449969 529.0490041 413.2959992 596.516698  442.6672517 387.5347163
 521.1953105 424.2028271 550.0475806 494.5518611 497.3895578 483.1597208
 442.7228916 581.7987977 565.9943634 368.6547849 505.1133435 319.9288698
 475.0154071 439.0747667 587.5747995 407.8763782 604.8413188 484.8769649
 515.8288149 452.3156755 560.5601606 505.7711403 512.8253581 426.775216
 402.1671222 578.9862586 524.7976276 532.9352188 630.1567282 553.6015347
 302.1895478 503.3878873 448.9332932 390.103273  434.144202  487.5475049
```

```
467.8009244 407.7045475 463.4849954 518.7864831 484.5198091 497.5866713
642.1015787 647.6194557 492.1050524 508.735741  479.6148117 532.7174857
563.4460357 479.2310929 506.5363931 275.9184207 576.3111774 514.0098178
424.7626355 444.5761441 571.2160048 475.2634237 529.2300901 612.3852299
593.1564015 557.5292736 516.8315567 560.4437922 426.1545477 409.0945262
482.8309859 508.7719067 630.4227632 542.7115581 491.0732237 408.9583359
536.4231045 507.4418323 448.2298292 614.7153338 536.7718994 507.212569
467.5019004 479.1728515 420.9161595 534.7771881 521.8835732 266.0863409
554.7220838 401.0331352 549.1315733 561.516532  528.9336186 486.8389348
408.2169018 527.7837898 586.1558702 502.1327892 461.7909591 476.1914133
536.1308969 410.0696111 535.4807752 488.3875258 533.3965538 529.0566632
414.9350607 478.8853913 357.8637186 568.7175759 493.1802162 375.3984554
469.3108615 700.9170916 387.3570727 604.3348401 431.6177338 492.9450531
497.81193   438.417742  557.6082621 532.7517876 446.4186734 461.6282784
475.7590678 528.309225  584.105885  573.4158673 432.4720613 512.1658664
510.6617922 458.3769107 357.5914394 338.3198626 551.6201455 444.5455497
459.2851235 465.8893127 555.8925954 457.8476959 476.1392469 549.0082269
532.1274491 439.9978799 376.3369008 640.18774   483.5431939 404.8245289
573.8474377 663.0748176 448.340425  550.8133677 605.0610388 422.3687366
506.42386   585.9318443 547.1109824 532.7248055 581.852344  486.1637991
541.226584  511.97986   626.0186727 544.7798637 577.273455  491.9115051
487.5554581 570.6300981 557.634109  583.0796357 493.719193  314.4385183
489.812488  441.8966315 660.4251843 522.3374046 447.3690272 487.379306
554.0030934 347.7769266 551.0230017 496.6507081 501.8744303 540.9957391
725.5848141 468.6684656 547.7099886 608.2718166 511.0387861 495.1759504
502.4097853 613.5993234 538.9419745 552.9403455 537.2150527 519.3409891
547.2443434 510.5394217 587.951054  590.5627196 708.9351849 530.7667187
507.3900618 443.197221  494.6871558 357.8579836 537.7731625 531.9615505
574.6548434 514.0889577 543.1326263 543.3401663 308.5277466 501.7492333
451.5756852 689.7876042 479.7319491 592.6884532 430.5888826 562.0820454
442.3631174 461.9208769 424.1854943 350.0582002 330.594446  517.1651356
640.5840619 573.3062223 397.4205841 456.4695101 376.4968407 559.199048
392.4973992 533.5149353 576.8025474 494.6386098 304.1355916 618.8459704
432.4811686 474.5323294 611.0000251 596.4301726 399.9838716 571.3074949
382.4161079 378.4735664 453.1695024 542.4124767 574.7472197 483.7965221
529.1945189 591.1971782 497.5136833 407.542168  436.5807403 640.7861664
423.4705332 478.1703341 581.3089329 451.4574469 663.8036933 445.7498412
548.2803202 503.2173931 529.5376653 452.1226251 521.2407802 490.2066
684.163431  591.7810894 443.9656268 447.6879065 588.7126055 468.4457372
485.9231305 424.7287739 444.582165  356.6155679 461.1122484 537.8252823
388.9405488 479.7319376 471.6028844 577.7360249 554.900783  495.2994425
357.7831107 478.584286  475.0716299 443.4418601 662.9610878 490.6004425
689.2356998 413.3717831 488.270298  567.4750105 418.1500811 498.6355985
```

```
print(X_test)
```

```
[[32.4914466  12.53035737 37.8752191   2.47613905]
 [32.7111193  12.32629139 36.67387836  3.35027929]
 [33.07773079 11.46698422 35.67572763  1.80922959]
 [32.21292383 11.73299146 35.63395395  4.33186303]
 [32.77049216 11.37176736 35.26149812  4.03438613]
 [30.73772037 12.63660605 36.21376309  3.35784684]
 [33.56647439 12.23565925 37.27757338  2.53204406]
 [32.24635    11.30555143 37.13312676  1.70738973]
 [34.56868085 11.37808709 38.30447119  3.78493211]
 [31.94539575 12.96576148 36.9663889   6.07665364]
 [32.68624509 12.63857212 36.09722093  4.29773748]
 [32.67294353 12.27605698 37.19279353  3.98247151]
 [35.86023651 11.73066139 36.88214908  3.41620998]
 [34.39432665 12.80775183 38.55103029  1.81007988]
 [31.86483255 13.44340599 36.87831537  2.36108695]
 [32.83789305 13.18518117 35.92159519  1.82359518]
 [34.20053941 12.66780888 37.48704925  3.70162229]
 [33.34450869 10.96980287 35.97457811  2.62762497]
 [36.13966249 12.05026723 36.95964319  3.86486074]
 [32.86532717 11.98441752 37.0443614   3.45238858]
 [34.96760989 13.9194944  37.95201319  5.06669686]
 [33.10035775 11.83211223 36.84149164  3.61223915]
 [32.92261076 11.56811634 36.90937821  2.47175067]
 [33.1556997  12.93155027 38.16643556  3.85447386]
 [34.55576799 12.17052542 39.13109673  3.66310549]
 [32.14906052 10.04731474 37.18144731  3.53508843]
 [33.50308726 12.8779837  37.44102134  1.55915194]
 [32.99059904 10.44123506 35.9389625   2.89507516]
 [33.79038721 11.94234087 38.06341359  4.0818027 ]
 [32.00475302 11.39520943 37.33281446  3.803365  ]
 [31.86274111 14.03986726 37.02226899  3.73822517]
 [32.40173183 12.08930957 38.3099079   3.87333757]
 [33.01479222 11.76117233 37.57016384  3.83416967]
 [32.71251233 11.72447386 37.1531516   3.308443  ]
 [31.44744649 10.1016322  38.04345265  4.23829619]
 [33.68093695 11.20156988 37.83544773  2.20881368]
 [34.72908017 11.96689808 36.54759628  2.95744876]
 [33.94311858 11.48419869 36.83936583  2.40245383]
```

```
[33.55165061 11.93689516 35.90025278  4.54333241]
[31.26064687 13.26676035 36.9711951   2.26725111]
[33.26463207 10.73213134 36.14579171  4.08656634]
[34.56455771 13.14655143 37.33544589  3.87687518]
[33.08529799 13.09353728 38.31564796  4.75036007]
[33.63662446 11.23650676 37.67502073  5.25470893]
[33.55520742 11.55182117 36.62883429  2.83794316]
[31.92627203 11.10946073 37.26895887  2.66403418]
[34.33075044 13.72245368 35.77311639  2.90900846]
[33.78520721 13.03951103 36.31272657  2.01819463]
[31.6739155  12.32914702 37.07437107  3.98246232]
[32.42569728 11.44890154 37.58019043  2.58696798]
[34.17951757 12.58154773 35.4442647   3.13706898]
[33.42874704 10.63676108 37.57883524  2.92639645]
[32.30274837 12.81539265 37.95780983  4.61542631]
[33.92462481 11.91141556 38.2747022   2.91003792]
[34.312167   11.81058676 37.41413357  2.47359612]
[29.53242897 10.9612984  37.42021558  4.04642316]
[31.12809005 13.27895623 37.38718053  4.62607529]
[32.52976873 11.7477317  36.93988205  0.80151572]
```

```
print(Y_test)
```

```
[449.0703194 482.6024673 374.2696745 513.1531119 502.7710746 461.7807422
 466.4211988 327.3779526 528.2238094 657.0199239 571.4710341 523.6339351
 546.5566669 467.4278485 439.8912805 445.0621855 553.9946736 403.7669021
 595.8038189 490.7386321 712.3963268 502.0925279 398.1634685 570.4517259
 549.8605905 392.9922559 419.9387748 407.6571788 521.1429518 463.7459811
 556.2981412 506.5473071 514.3365583 476.7667242 418.6027421 403.8195198
 506.1323424 436.7205559 558.9481124 421.3266313 478.2621264 593.915003
 632.1235881 591.4377356 444.2859075 392.2049334 558.4272572 447.1876443
 475.7250679 420.7376732 527.7829958 421.9667942 576.0252441 460.0612774
 452.627255  408.6403511 557.2526867 298.7620079 555.0683941 486.0834255
 488.7860611 423.3083341 458.7811317 574.4156896 489.9080531 352.5501082
 492.6060127 427.3565308 628.0478039 478.7193569 463.923513  765.5184619
 501.1224915 462.6565189 373.8857237 384.3260571 614.7296376 518.064558
 510.5014785 432.7207178 473.9498574 501.2091727 535.3216101 496.9334463
 442.0644138 505.2300683 544.4092722 744.2218671 392.810345  522.4041413
 519.3729768 528.4193297 385.0950071 504.8704324 501.1002452 385.152338
 520.8987945 499.1401524 616.851523  412.0129313 378.3309069 478.6009159
 595.8228367 492.5568337 521.5721748 444.0538266 422.4267759 501.9282649
 282.4712457 503.9783791 582.4919237 549.9041461 433.0487691 570.200409
 392.2852442 486.9470538 478.9514048 463.591418  496.5540816 524.6379646
 556.1863689 416.3583536 407.8040306 594.2744834 493.1812614]
```

```
sc=StandardScaler()
X_train=sc.fit_transform(X_train)
X_test =sc.transform(X_test)
```

```
print(X_train)
```

```
[[-1.43869753 -0.30628075  1.45477717  1.47980564]
 [ 0.72930762  0.00889083  0.1928528   0.40660492]
 [-0.21889638  0.01967932 -1.45632353 -1.19444246]
 ...
 [-0.01217758 -0.33299849 -1.0397505   1.46033234]
 [-0.36652251  0.94319928 -0.59122687  0.55196014]
 [ 1.5045378   1.2408939  -1.12869953  1.42357553]]
```

```
reg= LinearRegression()
reg.fit(X_train, Y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
reg.intercept_
```

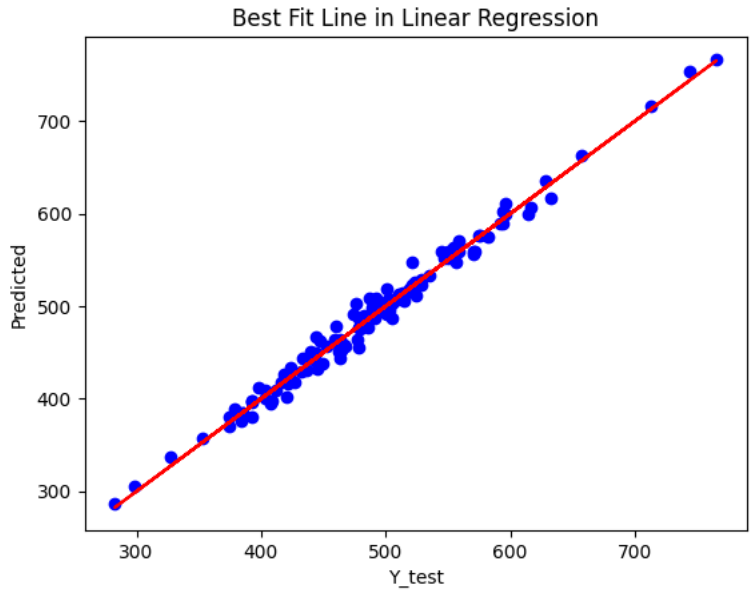
```
502.3944236152
```

```
reg.coef_
```

```
array([24.2517772 , 38.4022506 ,  0.2083891 , 60.60770351])
```

```
pred=reg.predict(X_test)
```

```
plt.scatter(Y_test, pred, color = 'blue')
plt.plot(Y_test,Y_test,color = 'red', linewidth = 1.5)
plt.xlabel('Y_test')
plt.ylabel('Predicted')
plt.title('Best Fit Line in Linear Regression')
plt.show()
```




```
sna.distplot((Y_test-pred),bins=50)
```

```
<ipython-input-92-57975588e45e>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sna.distplot((Y_test-pred),bins=50)
<Axes: ylabel='Density'>
```