ECON 427 HW 1 Yihua Zhang (905231562)

1.

(a)

These are the datasets for the first question.

@data

63.0278175,22.55258597,39.60911701,40.47523153,98.67291675,-0.254399986,Abnormal
39.05695098,10.06099147,25.01537822,28.99595951,114.4054254,4.564258645,Abnormal
68.83202098,22.21848205,50.09219357,46.61353893,105.9851355,-3.530317314,Abnormal
69.29700807,24.65287791,44.31123813,44.64413017,101.8684951,11.21152344,Abnormal
49.71285934,9.652074879,28.317406,40.06078446,108.1687249,7.918500615,Abnormal
40.25019968,13.92190658,25.1249496,26.32829311,130.3278713,2.230651729,Abnormal
53.43292815,15.86433612,37.16593387,37.56859203,120.5675233,5.988550702,Abnormal
45.36675362,10.75561143,29.03834896,34.61114218,117.2700675,-10.67587083,Abnormal
43.79019026,13.5337531,42.69081398,30.25643716,125.0028927,13.28901817,Abnormal
36.68635286,5.010884121,41.9487509,31.67546874,84.24141517,0.664437117,Abnormal
49.70660953,13.04097405,31.33450009,36.66563548,108.6482654,-7.825985755,Abnormal
31.23238734,17.71581923,15.5,13.51656811,120.0553988,0.499751446,Abnormal
48.91555137,19.96455616,40.26379358,28.95099521,119.321358,8.028894629,Abnormal
53.5721702,20.46082824,33.1,33.11134196,110.9666978,7.044802938,Abnormal
57.30022656,24.1888846,46.99999999,33.11134196,116.8065868,5.766946943,Abnormal
44.31890674,12.53799164,36.098763,31.78091509,124.1158358,5.415825143,Abnormal
63.83498162,20.36250706,54.55243367,43.47247456,112.3094915,-0.622526643,Abnormal
31.27601184,3.14466948,32.56299592,28.13134236,129.0114183,3.623020073,Abnormal
38.69791243,13.44474904,31,25.25316339,123.1592507,1.429185758,Abnormal
41.72996308,12.25407408,30.12258646,29.475889,116.5857056,-1.244402488,Abnormal
43.92283983,14.17795853,37.8325467,29.7448813,134.4610156,6.451647637,Abnormal
54.91944259,21.06233245,42.19999999,33.85711014,125.2127163,2.432561437,Abnormal
63.07361096,24.41380271,53.99999999,38.65980825,106.4243295,15.77969683,Abnormal
45.54078988,13.06959759,30.29832059,32.47119229,117.9808303,-4.987129618,Abnormal
36.12568347,22.75875277,29,13.3669307,115.5771163,-3.237562489,Abnormal
54.12492019,26.65048856,35.32974693,27.47443163,121.447011,1.571204816,Abnormal
26.14792141,10.75945357,14,15.38846783,125.2032956,-10.09310817,Abnormal
43.58096394,16.5088837,46.99999999,27.07208024,109.271634,8.992815727,Abnormal
44.5510115,21.93114655,26.78591597,22.61986495,111.0729197,2.652320636,Abnormal
66.87921138,24.89199889,49.27859673,41.9872125,113.4770183,-2.005891748,Abnormal
50.81926781,15.40221253,42.52893886,35.41705528,112.192804,10.86956554,Abnormal
46.39026008,11.07904664,32.13655345,35.31121344,98.77454633,6.386831648,Abnormal
44.93667457,17.44383762,27.78057555,27.49283695,117.9803245,5.569619587,Abnormal
38.66325708,12.98644139,39.99999999,25.67681568,124.914118,2.703008052,Abnormal
59.59554032,31.99824445,46.56025198,27.59729587,119.3303537,1.474285836,Abnormal
31.48421834,7.82622134,24.28481815,23.657997,113.8331446,4.393080498,Abnormal
32.09098679,6.989378081,35.99819848,25.10160871,132.264735,6.413427708,Abnormal
35.70345781,19.44325311,20.7,16.26020471,137.5406125,-0.263489651,Abnormal
55.84328595,28.84744756,47.69054322,26.99583839,123.3118449,2.812426855,Abnormal

52.41938511,19.01156052,35.87265953,33.40782459,116.5597709,1.694705102,Abnormal
35.49244617,11.7016723,15.59036345,23.79077387,106.9388517,-3.460357991,Abnormal
46.44207842,8.39503589,29.0372302,38.04704253,115.4814047,2.045475795,Abnormal
53.85479842,19.23064334,32.77905978,34.62415508,121.6709148,5.329843204,Abnormal
66.28539377,26.32784484,47.49999999,39.95754893,121.2196839,-0.799624469,Abnormal
56.03021778,16.2979149,62.27527456,39.73230287,114.0231172,-2.325683841,Abnormal
50.91244034,23.01516931,46.99999999,27.89727103,117.4222591,-2.526701511,Abnormal
48.332638,22.22778399,36.18199318,26.10485401,117.3846251,6.481709096,Abnormal
41.35250407,16.57736351,30.70619135,24.77514057,113.2666746,-4.497957556,Abnormal
40.55735663,17.97778407,34,22.57957256,121.0462458,-1.537383074,Abnormal
41.76773173,17.89940172,20.0308863,23.86833001,118.3633889,2.062962549,Abnormal
55.28585178,20.44011836,34,34.84573342,115.8770174,3.558372358,Abnormal
74.43359316,41.55733141,27.7,32.87626175,107.9493045,5.000088788,Abnormal
50.20966979,29.76012218,36.10400731,20.44954761,128.2925148,5.740614083,Abnormal
30.14993632,11.91744524,34,18.23249108,112.6841408,11.46322327,Abnormal
41.17167989,17.32120599,33.46940277,23.85047391,116.3778894,-9.569249858,Abnormal
47.65772963,13.27738491,36.67998541,34.38034472,98.24978071,6.273012173,Abnormal
43.34960621,7.467468964,28.06548279,35.88213725,112.7761866,5.753277458,Abnormal
46.85578065,15.35151393,38,31.50426672,116.2509174,1.662705589,Abnormal
43.20318499,19.66314572,35,23.54003927,124.8461088,-2.919075955,Abnormal
48.10923638,14.93072472,35.56468278,33.17851166,124.0564518,7.947904861,Abnormal
74.37767772,32.05310438,78.77201304,42.32457334,143.5606905,56.12590603,Abnormal
89.68056731,32.70443487,83.13073216,56.97613244,129.9554764,92.02727682,Abnormal
44.529051,9.433234213,51.99999999,35.09581679,134.7117723,29.10657504,Abnormal
77.69057712,21.38064464,64.42944191,56.30993248,114.818751,26.93184095,Abnormal
76.1472121,21.93618556,82.96150249,54.21102654,123.9320096,10.43197194,Abnormal
83.93300857,41.28630543,61.99999999,42.64670314,115.012334,26.58810016,Abnormal
78.49173027,22.1817978,59.99999999,56.30993248,118.5303266,27.38321314,Abnormal
75.64973136,19.33979889,64.14868477,56.30993248,95.9036288,69.55130292,Abnormal
72.07627839,18.94617604,50.99999999,53.13010236,114.2130126,1.01004051,Abnormal
58.59952852,-0.261499046,51.49999999,58.86102756,102.0428116,28.05969711,Abnormal
72.56070163,17.38519079,51.99999999,55.17551084,119.1937238,32.10853735,Abnormal
86.90079431,32.9281677,47.79434664,53.97262661,135.0753635,101.7190919,Abnormal
84.97413208,33.02117462,60.85987263,51.95295747,125.6595336,74.33340864,Abnormal
55.512212,20.09515673,43.99999999,35.41705528,122.648753,34.55294641,Abnormal
72.2223343,23.07771056,90.99999999,49.14462374,137.7366546,56.80409277,Abnormal
70.22145219,39.82272448,68.11840309,30.39872771,148.5255624,145.3781432,Abnormal
86.75360946,36.04301632,69.22104479,50.71059314,139.414504,110.8607824,Abnormal
58.78254775,7.667044186,53.33894082,51.11550357,98.50115697,51.58412476,Abnormal
67.41253785,17.44279712,60.14464036,49.96974073,111.12397,33.15764573,Abnormal
47.74467877,12.08935067,38.99999999,35.6553281,117.5120039,21.68240136,Abnormal
77.10657122,30.46999418,69.48062839,46.63657704,112.1516,70.75908308,Abnormal
74.00554124,21.12240192,57.37950226,52.88313932,120.2059626,74.55516588,Abnormal
88.62390839,29.08945331,47.56426247,59.53445508,121.7647796,51.80589921,Abnormal

81.10410039,24.79416792,77.88702048,56.30993247,151.8398566,65.21461611,Abnormal
76.32600187,42.39620445,57.19999999,33.92979742,124.267007,50.12745689,Abnormal
45.44374959,9.906071798,44.99999999,35.53767779,163.0710405,20.31531532,Abnormal
59.78526526,17.87932332,59.20646143,41.90594194,119.3191109,22.12386874,Abnormal
44.91414916,10.21899563,44.63091389,34.69515353,130.0756599,37.36453993,Abnormal
56.60577127,16.80020017,41.99999999,39.80557109,127.2945222,24.0185747,Abnormal
71.18681115,23.89620111,43.6966651,47.29061004,119.8649383,27.28398451,Abnormal
81.65603206,28.74886935,58.23282055,52.9071627,114.7698556,30.60914842,Abnormal
70.95272771,20.15993121,62.85910914,50.7927965,116.1779325,32.522331,Abnormal
85.35231529,15.84491006,71.66865979,69.50740523,124.4197875,76.0206034,Abnormal
58.10193455,14.83763914,79.64983825,43.26429541,113.5876551,50.23787808,Abnormal
94.17482232,15.38076983,67.70572132,78.79405249,114.8901128,53.25522004,Abnormal
57.52235608,33.64707522,50.90985841,23.87528085,140.9817119,148.7537109,Abnormal
96.65731511,19.46158117,90.21149828,77.19573393,120.6730408,64.08099841,Abnormal
74.72074622,19.75694203,82.73535954,54.96380419,109.3565941,33.30606685,Abnormal
77.65511874,22.4329501,93.89277881,55.22216863,123.0557067,61.2111866,Abnormal
58.52162283,13.92228609,41.46785522,44.59933674,115.514798,30.3879839,Abnormal
84.5856071,30.36168482,65.47948563,54.22392228,108.0102185,25.11847846,Abnormal
79.93857026,18.7740711,63.31183486,61.16449915,114.787107,38.53874133,Abnormal
70.39930842,13.46998624,61.19999999,56.92932218,102.3375244,25.53842852,Abnormal
49.78212054,6.46680486,52.99999999,43.31531568,110.8647831,25.33564729,Abnormal
77.40933294,29.39654543,63.23230243,48.0127875,118.4507311,93.56373734,Abnormal
65.00796426,27.60260762,50.94751899,37.40535663,116.5811088,7.015977884,Abnormal
65.01377322,9.838262375,57.73583722,55.17551084,94.73852542,49.69695462,Abnormal
78.42595126,33.42595126,76.27743927,45,138.5541111,77.15517241,Abnormal
63.17298709,6.330910974,62.99999999,56.84207612,110.6440206,42.60807567,Abnormal
68.61300092,15.0822353,63.01469619,53.53076561,123.4311742,39.49798659,Abnormal
63.90063261,13.7062037,62.12433389,50.19442891,114.1292425,41.42282844,Abnormal
84.99895554,29.61009772,83.35219438,55.38885782,126.9129899,71.32117542,Abnormal
42.02138603,-6.554948347,67.89999999,48.57633437,111.5857819,27.33867086,Abnormal
69.75666532,19.27929659,48.49999999,50.47736873,96.49136982,51.1696403,Abnormal
80.98807441,36.84317181,86.96060151,44.1449026,141.0881494,85.87215224,Abnormal
129.8340406,8.404475005,48.38405705,121.4295656,107.690466,418.5430821,Abnormal
70.48410444,12.48948765,62.41714208,57.99461679,114.1900488,56.90244779,Abnormal
86.04127982,38.75066978,47.87140494,47.29061004,122.0929536,61.98827709,Abnormal
65.53600255,24.15748726,45.77516991,41.3785153,136.4403015,16.37808564,Abnormal
60.7538935,15.7538935,43.19915768,45,113.0533309,31.69354839,Abnormal
54.74177518,12.09507205,40.99999999,42.64670314,117.6432188,40.3823266,Abnormal
83.87994081,23.07742686,87.14151223,60.80251395,124.6460723,80.55560527,Abnormal
80.07491418,48.06953097,52.40343873,32.00538321,110.7099121,67.72731595,Abnormal
65.66534698,10.54067533,56.48913545,55.12467166,109.1627768,53.93202006,Abnormal
74.71722805,14.32167879,32.5,60.39554926,107.1822176,37.01708012,Abnormal
48.06062649,5.687032126,57.05716117,42.37359436,95.44375749,32.83587702,Abnormal
70.67689818,21.70440224,59.18116082,48.97249594,103.0083545,27.8101478,Abnormal

80.43342782,16.998479,66.53601753,63.43494882,116.4389807,57.78125,Abnormal
90.51396072,28.27250132,69.8139423,62.2414594,100.8921596,58.82364821,Abnormal
77.23689752,16.73762214,49.77553438,60.49927538,110.6903772,39.7871542,Abnormal
50.06678595,9.120340183,32.16846267,40.94644577,99.71245318,26.76669655,Abnormal
69.78100617,13.77746531,57.99999999,56.00354085,118.9306656,17.91456046,Abnormal
69.62628302,21.12275138,52.76659472,48.50353164,116.8030913,54.81686729,Abnormal
81.75441933,20.12346562,70.56044038,61.63095371,119.4250857,55.50688907,Abnormal
52.20469309,17.21267289,78.09496877,34.9920202,136.9725168,54.93913416,Abnormal
77.12134424,30.3498745,77.48108264,46.77146974,110.6111484,82.09360704,Abnormal
88.0244989,39.84466878,81.77447308,48.17983012,116.6015376,56.76608323,Abnormal
83.39660609,34.31098931,78.42329287,49.08561678,110.4665164,49.67209559,Abnormal
72.05403412,24.70073725,79.87401586,47.35329687,107.1723576,56.42615873,Abnormal
85.09550254,21.06989651,91.73479193,64.02560604,109.062312,38.03283108,Abnormal
69.56348614,15.4011391,74.43849743,54.16234705,105.0673556,29.70121083,Abnormal
89.5049473,48.90365265,72.0034229,40.60129465,134.6342912,118.3533701,Abnormal
85.29017283,18.27888963,100.7442198,67.0112832,110.6607005,58.88494802,Abnormal
60.62621697,20.5959577,64.53526221,40.03025927,117.2255542,104.8592474,Abnormal
60.04417717,14.30965614,58.03886519,45.73452103,105.1316639,30.40913315,Abnormal
85.64378664,42.68919513,78.7506635,42.95459151,105.1440758,42.88742577,Abnormal
85.58171024,30.45703858,78.23137949,55.12467166,114.8660487,68.37612182,Abnormal
55.08076562,-3.759929872,55.99999999,58.84069549,109.9153669,31.77358318,Abnormal
65.75567895,9.832874231,50.82289501,55.92280472,104.3949585,39.30721246,Abnormal
79.24967118,23.94482471,40.79669829,55.30484647,98.62251165,36.7063954,Abnormal
81.11260488,20.69044356,60.68700588,60.42216132,94.01878339,40.51098228,Abnormal
48.0306238,3.969814743,58.34451924,44.06080905,125.3509625,35.00007784,Abnormal
63.40448058,14.11532726,48.13680562,49.28915333,111.9160075,31.78449499,Abnormal
57.28694488,15.1493501,63.99999999,42.13759477,116.7353868,30.34120327,Abnormal
41.18776972,5.792973871,42.86739151,35.39479584,103.3488802,27.66027669,Abnormal
66.80479632,14.55160171,72.08491177,52.25319461,82.45603817,41.6854736,Abnormal
79.4769781,26.73226755,70.65098189,52.74471055,118.5886691,61.70059824,Abnormal
44.21646446,1.507074501,46.11033909,42.70938996,108.6295666,42.81048066,Abnormal
57.03509717,0.34572799,49.19800263,56.68936918,103.0486975,52.16514503,Abnormal
64.27481758,12.50864276,68.70237672,51.76617482,95.25245421,39.40982612,Abnormal
92.02630795,35.39267395,77.41696348,56.633634,115.72353,58.05754155,Abnormal
67.26314926,7.194661096,51.69688681,60.06848816,97.8010854,42.13694325,Abnormal
118.1446548,38.44950127,50.83851954,79.69515353,81.0245406,74.04376736,Abnormal
115.9232606,37.51543601,76.79999999,78.40782459,104.6986033,81.19892712,Abnormal
53.94165809,9.306594428,43.10049819,44.63506366,124.3978211,25.0821266,Abnormal
83.7031774,20.26822858,77.1105979,63.43494882,125.4801739,69.279571,Abnormal
56.99140382,6.87408897,57.00900516,50.11731485,109.978045,36.81011057,Abnormal
72.34359434,16.42078962,59.86901238,55.92280472,70.08257486,12.07264427,Abnormal
95.38259648,24.82263131,95.15763273,70.55996517,89.3075466,57.66084135,Abnormal
44.25347645,1.101086714,38,43.15238973,98.27410705,23.9106354,Abnormal
64.80954139,15.17407796,58.83999352,49.63546343,111.679961,21.40719845,Abnormal

78.40125389,14.04225971,79.69426258,64.35899418,104.7312342,12.39285327,Abnormal
56.66829282,13.45820343,43.76970978,43.21008939,93.69220863,21.10812135,Abnormal
50.82502875,9.064729049,56.29999999,41.7602997,78.99945411,23.04152435,Abnormal
61.41173702,25.38436364,39.09686927,36.02737339,103.4045971,21.84340688,Abnormal
56.56382381,8.961261611,52.57784639,47.6025622,98.77711506,50.70187326,Abnormal
67.02766447,13.28150221,66.15040334,53.74616226,100.7154129,33.98913551,Abnormal
80.81777144,19.23898066,61.64245116,61.57879078,89.47183446,44.167602,Abnormal
80.65431956,26.34437939,60.89811835,54.30994017,120.1034928,52.46755185,Abnormal
68.72190982,49.4318636,68.0560124,19.29004622,125.0185168,54.69128928,Abnormal
37.90391014,4.47909896,24.71027447,33.42481118,157.848799,33.60702661,Abnormal
64.62400798,15.22530262,67.63216653,49.39870535,90.298468,31.32641123,Abnormal
75.43774787,31.53945399,89.59999999,43.89829388,106.8295898,54.96578902,Abnormal
71.00194076,37.51577195,84.53709256,33.48616882,125.1642324,67.77118983,Abnormal
81.05661087,20.80149217,91.78449512,60.2551187,125.430176,38.18178176,Abnormal
91.46874146,24.50817744,84.62027202,66.96056402,117.3078968,52.62304673,Abnormal
81.08232025,21.25584028,78.76675639,59.82647997,90.07187999,49.159426,Abnormal
60.419932,5.265665422,59.8142356,55.15426658,109.0330745,30.26578534,Abnormal
85.68094951,38.65003527,82.68097744,47.03091424,120.8407069,61.95903428,Abnormal
82.4065243,29.27642195,77.05456489,53.13010235,117.0422439,62.76534831,Abnormal
43.7182623,9.811985315,51.99999999,33.90627699,88.43424213,40.88092253,Abnormal
86.472905,40.30376567,61.14101155,46.16913933,97.4041888,55.75222146,Abnormal
74.46908181,33.28315665,66.94210105,41.18592517,146.4660009,124.9844057,Abnormal
70.25043628,10.34012252,76.37007032,59.91031376,119.2370072,32.66650243,Abnormal
72.64385013,18.92911726,67.99999999,53.71473287,116.9634162,25.38424676,Abnormal
71.24176388,5.268270454,85.99958417,65.97349342,110.703107,38.2598637,Abnormal
63.7723908,12.76338484,65.36052425,51.00900596,89.82274067,55.99545386,Abnormal
58.82837872,37.57787321,125.7423855,21.25050551,135.6294176,117.3146829,Abnormal
74.85448008,13.90908417,62.69325884,60.9453959,115.2087008,33.17225512,Abnormal
75.29847847,16.67148361,61.29620362,58.62699486,118.8833881,31.57582292,Abnormal
63.36433898,20.02462134,67.49870507,43.33971763,130.9992576,37.55670552,Abnormal
67.51305267,33.2755899,96.28306169,34.23746278,145.6010328,88.30148594,Abnormal
76.31402766,41.93368293,93.2848628,34.38034472,132.2672855,101.2187828,Abnormal
73.63596236,9.711317947,62.99999999,63.92464442,98.72792982,26.97578722,Abnormal
56.53505139,14.37718927,44.99154663,42.15786212,101.7233343,25.77317356,Abnormal
80.11157156,33.94243223,85.10160773,46.16913933,125.5936237,100.2921068,Abnormal
95.48022873,46.55005318,58.99999999,48.93017555,96.68390337,77.28307195,Abnormal
74.09473084,18.82372712,76.03215571,55.27100372,128.4057314,73.38821617,Abnormal
87.67908663,20.36561331,93.82241589,67.31347333,120.9448288,76.73062904,Abnormal
48.25991962,16.41746236,36.32913708,31.84245726,94.88233607,28.34379914,Abnormal
38.50527283,16.96429691,35.11281407,21.54097592,127.6328747,7.986683227,Normal
54.92085752,18.96842952,51.60145541,35.952428,125.8466462,2.001642472,Normal
44.36249017,8.945434892,46.90209626,35.41705528,129.220682,4.994195288,Normal
48.3189305,17.45212105,47.99999999,30.86680945,128.9803079,-0.910940567,Normal
45.70178875,10.65985935,42.5778464,35.0419294,130.1783144,-3.38890999,Normal

30.74193812,13.35496594,35.90352597,17.38697218,142.4101072,-2.005372903,Normal
50.91310144,6.6769999,30.89652243,44.23610154,118.151531,-1.057985526,Normal
38.12658854,6.557617408,50.44507473,31.56897113,132.114805,6.338199339,Normal
51.62467183,15.96934373,35,35.6553281,129.385308,1.00922834,Normal
64.31186727,26.32836901,50.95896417,37.98349826,106.1777511,3.118221289,Normal
44.48927476,21.78643263,31.47415392,22.70284212,113.7784936,-0.284129366,Normal
54.9509702,5.865353416,52.99999999,49.08561678,126.9703283,-0.631602951,Normal
56.10377352,13.10630665,62.63701952,42.99746687,116.2285032,31.17276727,Normal
69.3988184,18.89840693,75.96636144,50.50041147,103.5825398,-0.44366081,Normal
89.83467631,22.63921678,90.56346144,67.19545953,100.5011917,3.040973261,Normal
59.72614016,7.724872599,55.34348527,52.00126756,125.1742214,3.235159224,Normal
63.95952166,16.06094486,63.12373633,47.8985768,142.3601245,6.298970934,Normal
61.54059876,19.67695713,52.89222856,41.86364163,118.6862678,4.815031084,Normal
38.04655072,8.30166942,26.23683004,29.7448813,123.8034132,3.885773488,Normal
43.43645061,10.09574326,36.03222439,33.34070735,137.4396942,-3.114450861,Normal
65.61180231,23.13791922,62.58217893,42.47388309,124.1280012,-4.083298414,Normal
53.91105429,12.93931796,38.99999999,40.97173633,118.1930354,5.074353176,Normal
43.11795103,13.81574355,40.34738779,29.30220748,128.5177217,0.970926407,Normal
40.6832291,9.148437195,31.02159252,31.53479191,139.1184721,-2.511618596,Normal
37.7319919,9.386298276,41.99999999,28.34569362,135.740926,13.68304672,Normal
63.92947003,19.97109671,40.17704963,43.95837332,113.0659387,-11.05817866,Normal
61.82162717,13.59710457,63.99999999,48.22452261,121.779803,1.296191194,Normal
62.14080535,13.96097523,57.99999999,48.17983012,133.2818339,4.955105669,Normal
69.00491277,13.29178975,55.5701429,55.71312302,126.6116215,10.83201105,Normal
56.44702568,19.44449915,43.5778464,37.00252653,139.1896903,-1.859688529,Normal
41.6469159,8.835549101,36.03197484,32.8113668,116.5551679,-6.054537956,Normal
51.52935759,13.51784732,35,38.01151027,126.7185156,13.92833085,Normal
39.08726449,5.536602477,26.93203835,33.55066201,131.5844199,-0.75946135,Normal
34.64992241,7.514782784,42.99999999,27.13513962,123.9877408,-4.082937601,Normal
63.02630005,27.33624023,51.60501665,35.69005983,114.5066078,7.439869802,Normal
47.80555887,10.68869819,53.99999999,37.11686068,125.3911378,-0.402523218,Normal
46.63786363,15.85371711,39.99999999,30.78414653,119.3776026,9.06458168,Normal
49.82813487,16.73643493,28,33.09169994,121.4355585,1.91330704,Normal
47.31964755,8.573680295,35.56025198,38.74596726,120.5769719,1.630663508,Normal
50.75329025,20.23505957,37,30.51823068,122.343516,2.288487746,Normal
36.15782981,-0.810514093,33.62731353,36.96834391,135.9369096,-2.092506504,Normal
40.74699612,1.835524271,49.99999999,38.91147185,139.2471502,0.668556793,Normal
42.91804052,-5.845994341,57.99999999,48.76403486,121.6068586,-3.362044654,Normal
63.79242525,21.34532339,65.99999999,42.44710185,119.5503909,12.38260373,Normal
72.95564397,19.57697146,61.00707117,53.37867251,111.2340468,0.813491154,Normal
67.53818154,14.65504222,58.00142908,52.88313932,123.6322597,25.9702063,Normal
54.75251965,9.752519649,47.99999999,45,123.0379985,8.235294118,Normal
50.16007802,-2.970024337,41.99999999,53.13010235,131.8024914,-8.290203373,Normal
40.34929637,10.19474845,37.96774659,30.15454792,128.0099272,0.458901373,Normal

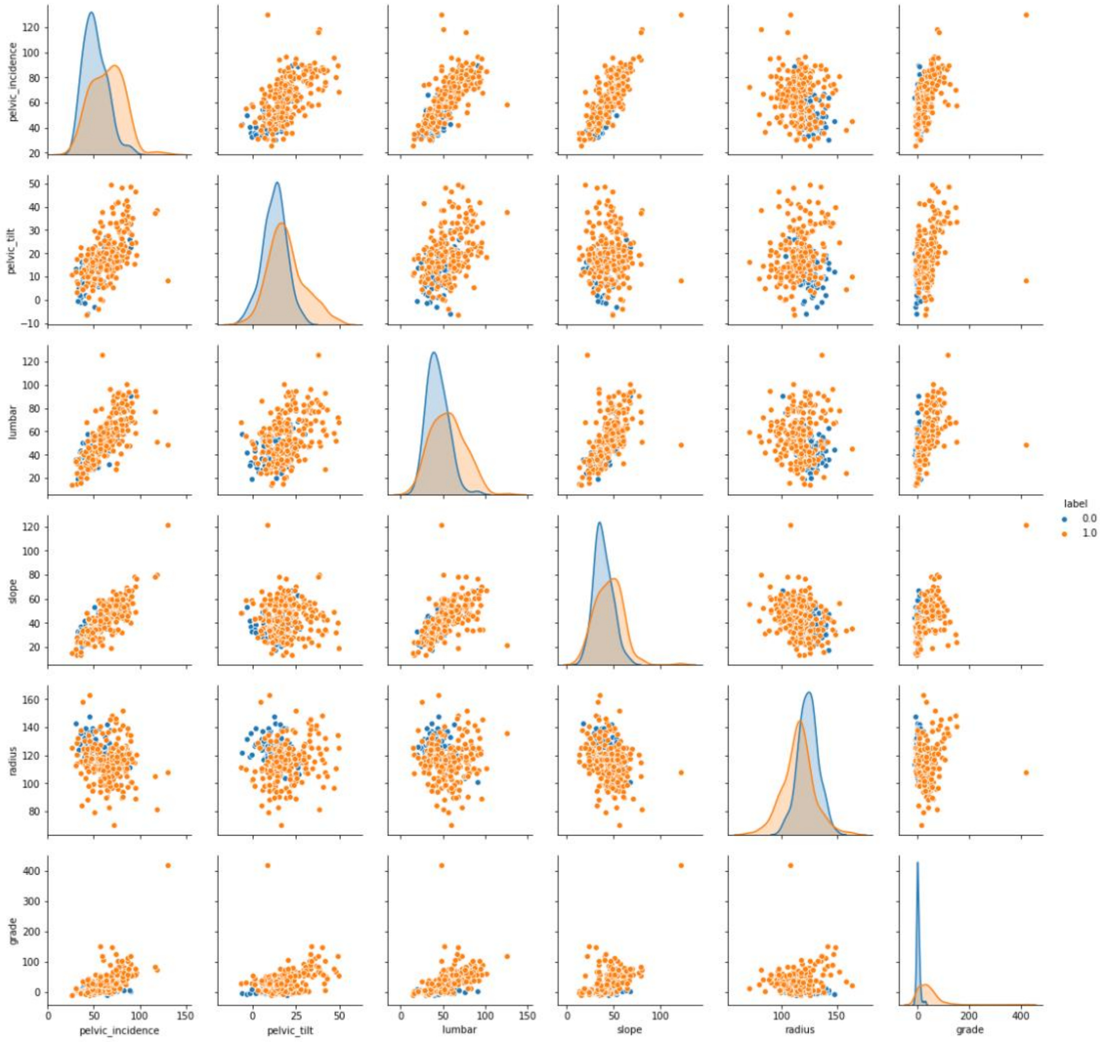
63.61919213,16.93450781,49.34926218,46.68468432,117.0897469,-0.357811974,Normal
54.14240778,11.93511014,42.99999999,42.20729763,122.2090834,0.153549242,Normal
74.97602148,14.92170492,53.73007172,60.05431656,105.6453997,1.594747729,Normal
42.51727249,14.37567126,25.32356538,28.14160123,128.9056892,0.75702014,Normal
33.78884314,3.675109986,25.5,30.11373315,128.3253556,-1.776111234,Normal
54.5036853,6.819910138,46.99999999,47.68377516,111.7911722,-4.406769011,Normal
48.17074627,9.594216702,39.71092029,38.57652956,135.6233101,5.360050572,Normal
46.37408781,10.21590237,42.69999999,36.15818544,121.2476572,-0.54202201,Normal
52.86221391,9.410371613,46.98805181,43.4518423,123.0912395,1.856659161,Normal
57.1458515,16.48909145,42.84214764,40.65676005,113.8061775,5.0151857,Normal
37.14014978,16.48123972,24,20.65891006,125.0143609,7.366425398,Normal
51.31177106,8.875541276,56.99999999,42.43622979,126.4722584,-2.144043911,Normal
42.51561014,16.54121618,41.99999999,25.97439396,120.631941,7.876730692,Normal
39.35870531,7.011261806,37,32.3474435,117.8187599,1.904048199,Normal
35.8775708,1.112373561,43.45725694,34.76519724,126.9239062,-1.632238263,Normal
43.1919153,9.976663803,28.93814927,33.21525149,123.4674001,1.741017579,Normal
67.28971201,16.7175142,50.99999999,50.5721978,137.5917777,4.960343813,Normal
51.32546366,13.63122319,33.25857782,37.69424047,131.3061224,1.78886965,Normal
65.7563482,13.20692644,43.99999999,52.54942177,129.3935728,-1.982120038,Normal
40.41336566,-1.329412398,30.98276809,41.74277806,119.3356546,-6.173674823,Normal
48.80190855,18.01776202,51.99999999,30.78414653,139.1504066,10.44286169,Normal
50.08615264,13.43004422,34.45754051,36.65610842,119.1346221,3.089484465,Normal
64.26150724,14.49786554,43.90250363,49.76364169,115.3882683,5.951454368,Normal
53.68337998,13.44702168,41.58429713,40.23635831,113.9137026,2.737035292,Normal
48.99595771,13.11382047,51.87351997,35.88213725,126.3981876,0.535471617,Normal
59.16761171,14.56274875,43.19915768,44.60486296,121.0356423,2.830504124,Normal
67.80469442,16.55066167,43.25680184,51.25403274,119.6856451,4.867539941,Normal
61.73487533,17.11431203,46.89999999,44.6205633,120.9201997,3.087725997,Normal
33.04168754,-0.324678459,19.0710746,33.366366,120.3886112,9.354364925,Normal
74.56501543,15.72431994,58.61858244,58.84069549,105.417304,0.599247113,Normal
44.43070103,14.17426387,32.2434952,30.25643716,131.7176127,-3.604255336,Normal
36.42248549,13.87942449,20.24256187,22.543061,126.0768612,0.179717077,Normal
51.07983294,14.20993529,35.95122893,36.86989765,115.8037111,6.905089963,Normal
34.75673809,2.631739646,29.50438112,32.12499844,127.1398495,-0.460894198,Normal
48.90290434,5.587588658,55.49999999,43.31531568,137.1082886,19.85475919,Normal
46.23639915,10.0627701,37,36.17362905,128.0636203,-5.100053328,Normal
46.42636614,6.620795049,48.09999999,39.80557109,130.3500956,2.449382401,Normal
39.65690201,16.20883944,36.67485694,23.44806258,131.922009,-4.968979881,Normal
45.57548229,18.75913544,33.77414297,26.81634684,116.7970069,3.131909921,Normal
66.50717865,20.89767207,31.72747138,45.60950658,128.9029049,1.517203356,Normal
82.90535054,29.89411893,58.25054221,53.01123161,110.7089577,6.079337831,Normal
50.67667667,6.461501271,35,44.2151754,116.5879699,-0.214710615,Normal
89.01487529,26.07598143,69.02125897,62.93889386,111.4810746,6.061508401,Normal
54.60031622,21.48897426,29.36021618,33.11134196,118.3433212,-1.471067262,Normal

34.38229939,2.062682882,32.39081996,32.31961651,128.3001991,-3.365515555,Normal
45.07545026,12.30695118,44.58317718,32.76849908,147.8946372,-8.941709421,Normal
47.90356517,13.61668819,36,34.28687698,117.4490622,-4.245395422,Normal
53.93674778,20.72149628,29.22053381,33.21525149,114.365845,-0.421010392,Normal
61.44659663,22.6949683,46.17034732,38.75162833,125.6707246,-2.707879517,Normal
45.25279209,8.693157364,41.5831264,36.55963472,118.5458418,0.214750167,Normal
33.84164075,5.073991409,36.64123294,28.76764934,123.9452436,-0.199249089,Normal

(b)
Here are the codes and graphs of (i), (ii) and (iii).
trykNN.py

```
##
import pandas as pd
import numpy as np
import math
import seaborn as sns
from sklearn.model_selection import train_test_split

users = pd.read_table('vertebral_column_data/column_2C.dat',sep = '::',header=None,engine='python')
users=users[0]
#loading the data and get into the list
datalist=np.zeros((len(users),7))
for idx in range (len(users)):
    Str=users[idx]
    Str=Str.split()

    for i in range(len(Str)):
        element=Str[i]
        if i<(len(Str)-1):
            datalist[idx][i]=float(element)
        else:
            if element=='AB':
                datalist[idx][i]=1
            else:
                datalist[idx][i]=0

newdf=pd.DataFrame({"label":datalist[:,6],"pelvic_incidence":datalist[:,0],"pelvic_tilt":datalist[:,1],"lumbar":datalist[:,2],"slope":datalist[:,3],"radius":datalist[:,4],"grade":datalist[:,5]})
##corresponding to b
sns.pairplot(data=newdf, hue='label', vars=['pelvic_incidence', 'pelvic_tilt', 'lumbar', 'slope','radius','grade'])
###
```

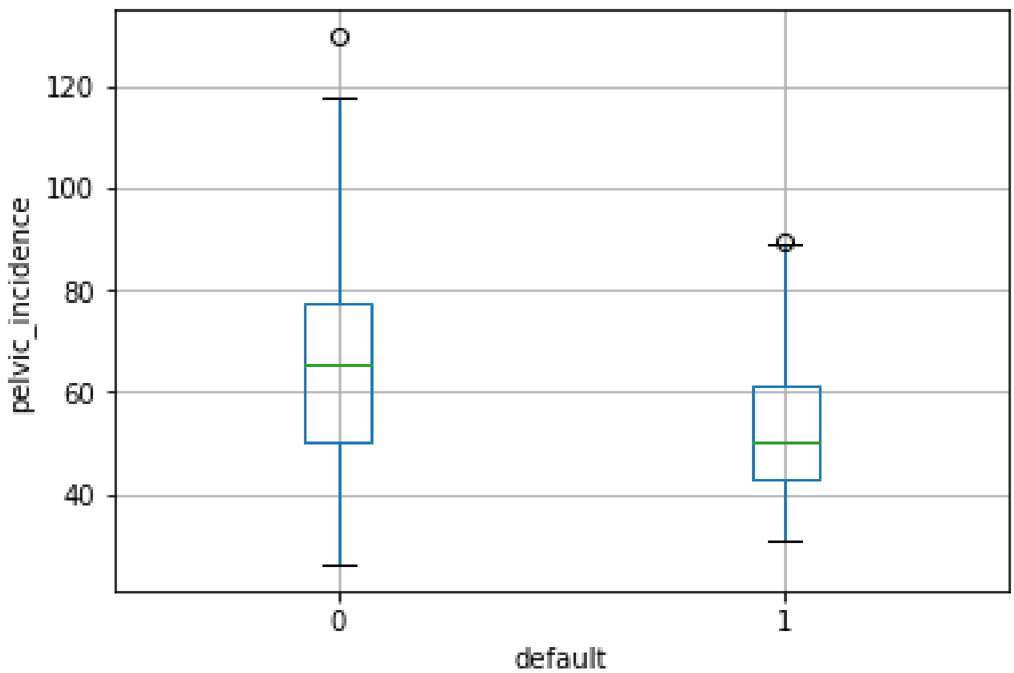
Here are the graphs for the questions:

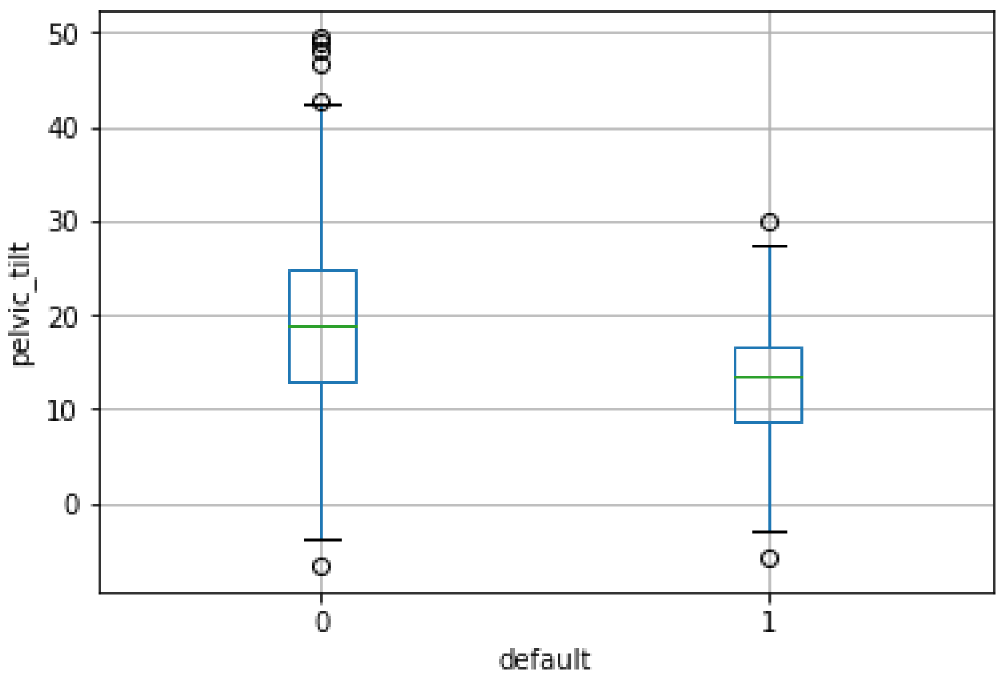


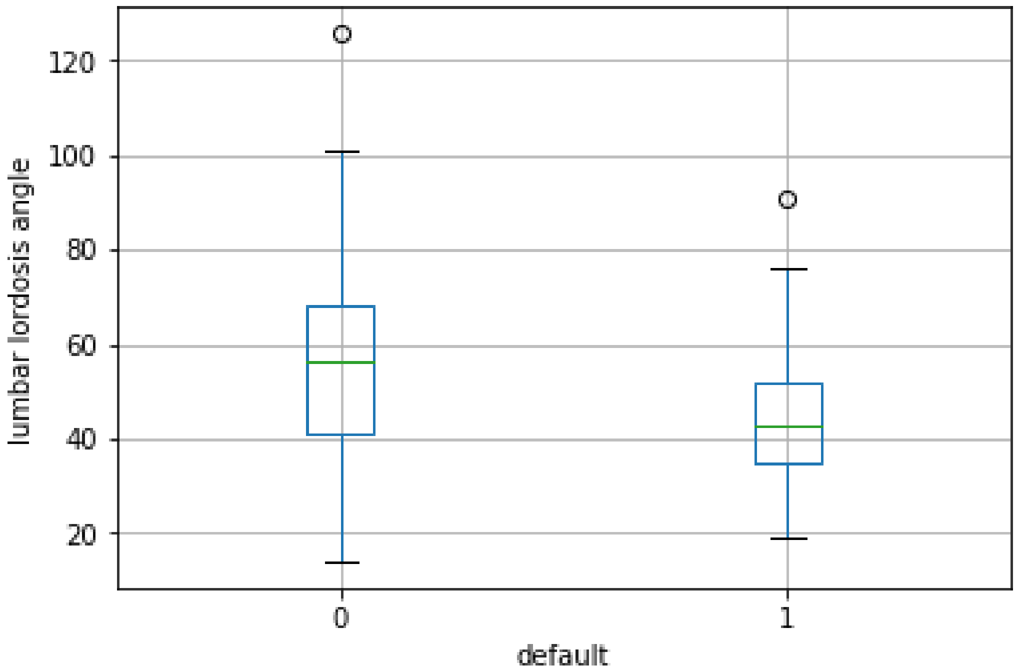
ii

Make boxplots: the following code is about pelvic incidence variables, the others boxplots are made in a similar way.

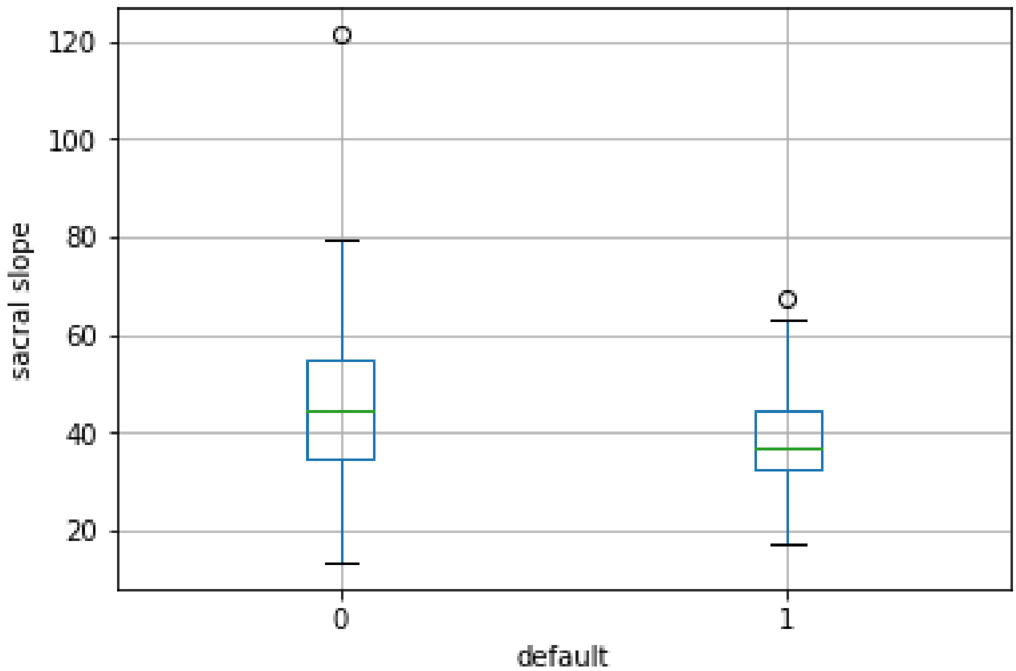
```
##
pel_1=newdf_1['pelvic_incidence']
pel_1=pel_1.rename(columns={'pelvic_incidence':'1'},inplace=True)
pel_0=newdf_0['pelvic_incidence']
pel_0=pel_0.rename(columns={'pelvic_incidence':'0'},inplace=True)

a=pd.concat([pel_1,pel_0],axis=1)
#plt.boxplot([pel_1,pel_0],patch_artist=True, boxprops={'color'})
a.boxplot()
plt.ylabel('pelvic_incidence')
plt.xlabel('default')
plt.show()
##
```


120
100
80
60
40
pelvic_incidence
0
1
default

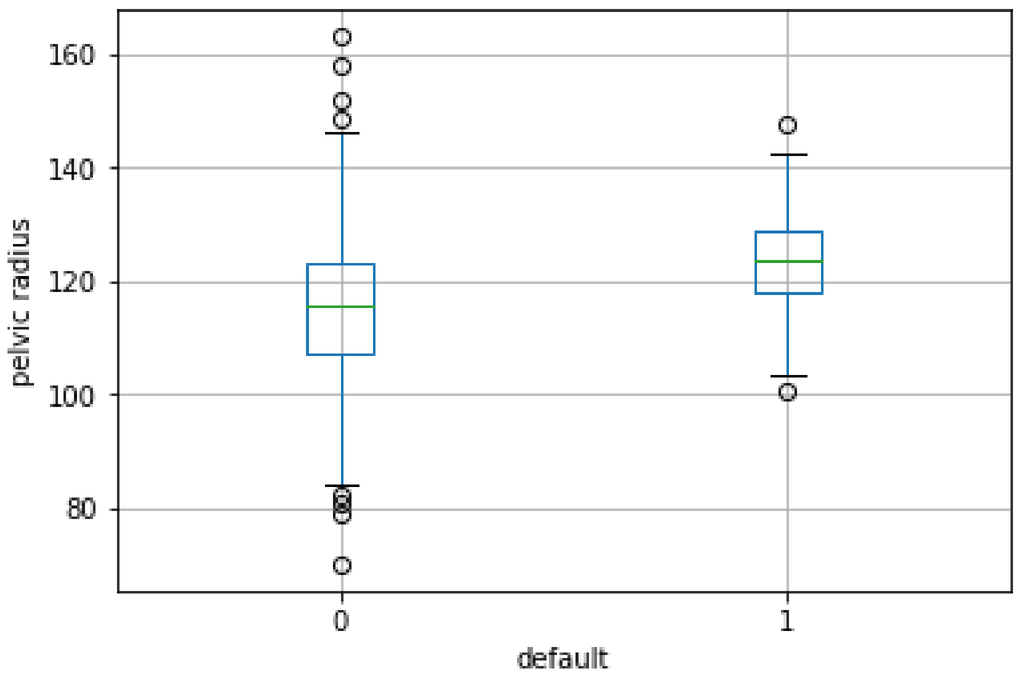

50
40
30
20
10
0
pelvic_tilt
0
1
default

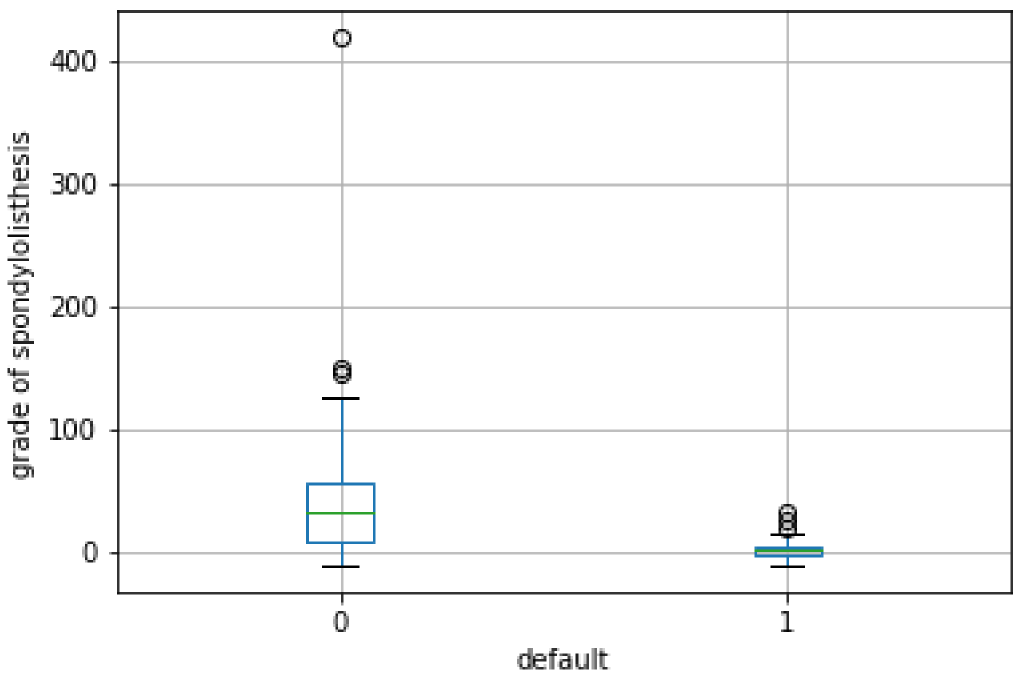


120
100
80
60
40
20
lumbar lordosis angle
0
1
default



120
100
80
60
40
20
sacral slope
0
1
default





iii

```
##
X=newdf[['pelvic_incidence','pelvic_tilt','lumbar','slope','radius','grade']]
y=newdf['label']

x0=X[y==0]
```

```
x1=X[y==1]

x0_train=x0.head(70)
y0_train=y[y==0].head(70)

x0_test=x0[70:len(x0)]
y0_test=y[y==0].head(x0_test.shape[0])

x1_train=x1.head(140)
y1_train=y[y==1].head(140)

x1_test=x1[140:len(x1)]
y1_test=y[y==1].head(x1_test.shape[0])

x_train=x0_train.append(x1_train)
y_train=y0_train.append(y1_train)

x_test=x0_test.append(x1_test)
y_test=y0_test.append(y1_test)

##
```

(c)

(i)

Here are the codes for the k-nearest neighbors with the Euclidean metric.

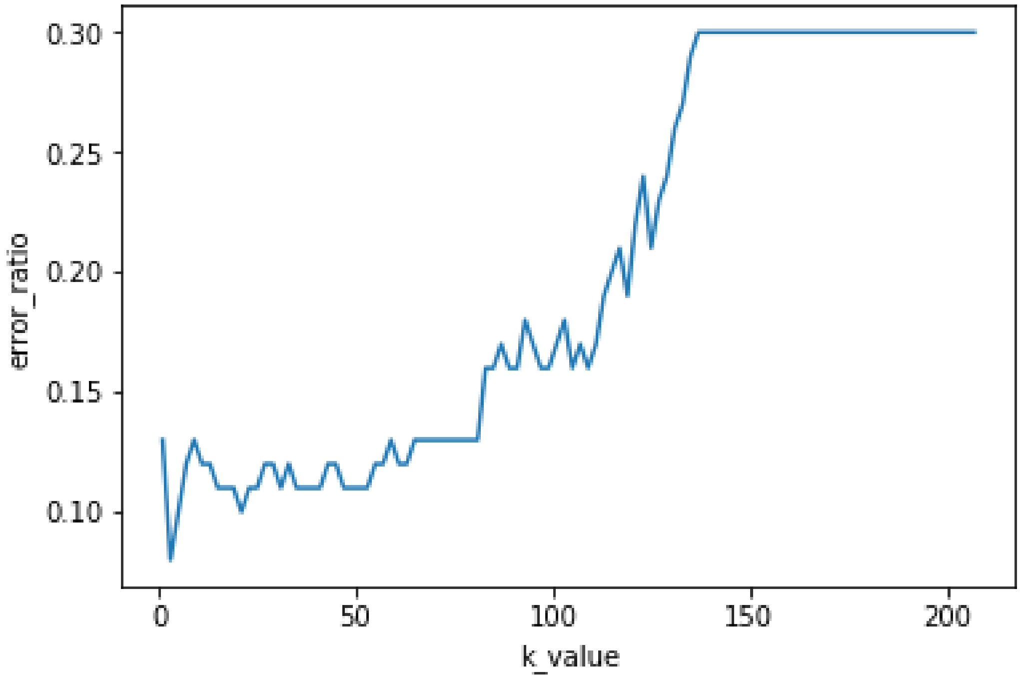
```
##
def Euclidean(single_data,norm_x_train):
    diffmat=np.tile(single_data,(norm_x_train.shape[0],1))-norm_x_train
    sq_diff=diffmat**2
    distance=(sq_diff.sum(axis=1))**0.5

    return distance    #Euclidean metric
##
```

(ii)

In this question…

trykNN3.py

k values：{1， 3， 5， ……209}



When K=3, the minimum error rate is 0.08.

Calculating the confusion matrix

```
##
tp=0
fn=0
fp=0
tn=0
for row in range (test_num):
    compare=best_compare[row,:]
    y_real=compare[0]
    y_pre=compare[1]
    if y_real==1 and y_pre==1:
        tp+=1
    else:
        if y_real==1 and y_pre==0:
            fn+=1
        else:
            if y_real==0 and y_pre==1:
                fp+=1
            else:
                tn+=1


confusion_matrix=[[tp,fp,tp+fp],[fn,tn,fn+tn],[tp+fn,fp+tn,tp+fn+fp+tn]]
##
```

Now, we will get the results like this…

Confusion matrix

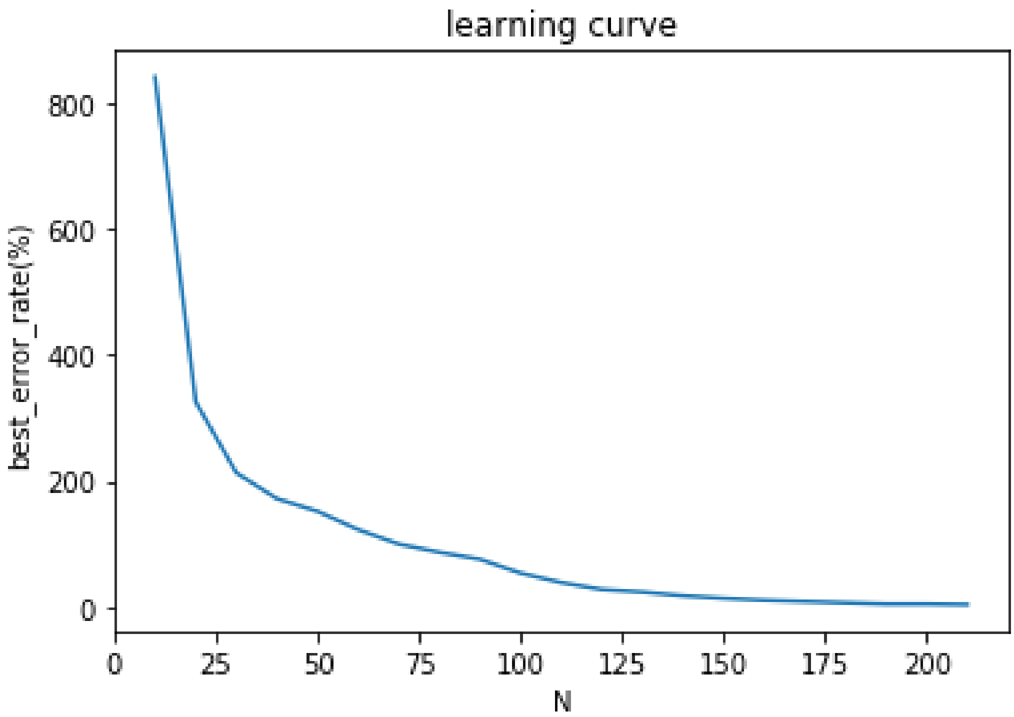
		True status		
		1	0	total
Predicted status	1	69	7	76
	0	1	23	24
	total	70	30	100

Precision=0.908

Fscore=0.945

(iii)

learning_curve.py



The codes are here：

```
##
N_range=np.arange(10,220,10)
N_best_error_ratio=np.zeros_like(N_range)

for N in N_range:
    x_train,y_train,x_test,y_test=prepare_data(x0,x1,y,N)
    norm_x_train=norm2(x_train)
    norm_x_test=norm2(x_test)

    test_num=norm_x_test.shape[0]
    y_train_dist=np.zeros((norm_x_train.shape[0],2))

    compare=np.zeros((test_num,2))
```

```
    k_value=np.arange(1,min(N,196+5),5)
    error_ratio=[]
    for k in k_value:
        error_sum=0

compare_result,error_sum_result=kNN_compare(k,norm_x_test,norm_x_train,y_test,test_num)
        error=error_sum_result/test_num
        error_ratio.append(error)

    min_index=np.array(error_ratio).argsort()
    k_best=k_value[min_index[0]]

best_compare,best_error=kNN_compare(k_best,norm_x_test,norm_x_train,y_test,test_num)
    idx_e=int(N/10-1)
    N_best_error_ratio[idx_e]=best_error/N*100



plt.figure(1)
plt.xlabel('N')
plt.ylabel('best_error_rate(%)')
plt.title('learning curve')
plt.plot(N_range,N_best_error_ratio)
##
```

(d)

(I)

The definition of Minkovski distance:

```
#
def Minkow(single_data,norm_x_train,pvalue):
    distance=np.zeros((norm_x_train.shape[0]))
    x1=single_data
    for row in range(norm_x_train.shape[0]):
        x2=norm_x_train[row,:]
        dist=DistanceMetric.get_metric('minkowski',p=pvalue)
        xmat=[x1,x2]
        distance[row]=dist.pairwise(xmat)[0][1]
    return distance
#
```

A. When p=1, it will be the Manhattan Distance.

The file is knn_d.py

When using Minkow function, make the p-value to 1.

In the kNN_compare function, the minimum value of the error rate in distance=Minkow(single_data,norm_x_train,1)
Is 0.11， the k is 6

B． For the（list）q_value，with log10(p)
The corresponding file is knn_d2.py

```
# q_value=np.arange(0.1,1.1,0.1)
##
for q in q_value:
    pvalue=10**q
    error_sum=0

compare_result,error_sum_result=kNN_compare_p(k,norm_x_test,norm_x_train,y_test,
test_num,pvalue)
    error=error_sum_result/test_num
    error_ratio.append(error)
##
distance=Minkow(single_data,norm_x_train,pvalue)
```

After the calculation,
The lowest error rate is 0.1, and for the corresponding log10(p) is 0.6.

C．

```
# distance=chebdist(single_data,norm_x_train)
##
def chebdist(single_data,norm_x_train):
    distance=np.zeros((norm_x_train.shape[0]))
    x1=single_data
    for row in range(norm_x_train.shape[0]):
        x2=norm_x_train[row,:]
        dist=DistanceMetric.get_metric('chebyshev')
        xmat=[x1,x2]
        distance[row]=dist.pairwise(xmat)[0][1]
    return distance
##
```

As the result, the smallest error rate is 0.1 and the k will be 16.

(ii)
Mahalanobis distance

```
#distance=Minkow(single_data,norm_x_train,pvalue)
def mahadist(single_data,norm_x_train):
    distance=np.zeros((norm_x_train.shape[0]))
    x1=single_data
    for row in range(norm_x_train.shape[0]):
        x2=norm_x_train[row,:]
```

```
X=np.vstack([x1,x2])
XT=X.T
S=np.cov(X)
SI=np.linalg.inv(S)
n=XT.shape[0]

for i in range(0,n):
    for j in range(i+1,n):
        delta=XT[i]-XT[j]
        d=np.sqrt(np.dot(np.dot(delta,SI),delta.T))
        distance[row]=d
```

For the return distance, the smallest error rate is 0.1 and the k will be 21.

(e)

```
##
        count_class={}
        for ik in range (k):
            vote_y=y_train.iloc[sort_dist[ik]]

            count_class[vote_y]=count_class.get(vote_y,0)+float(1/distance[sort_dist[ik]]) #

##
```

	Euclidean	Manhattan	Chebyshev
Majority vote	K=16 Error rate=0.1	K=6 Error rate=0.11	K=16 Error rate=0.07
Weight	K=6 Error rate=0.1	K=26 Error rate=0.1	K=16 Error rate=0.1