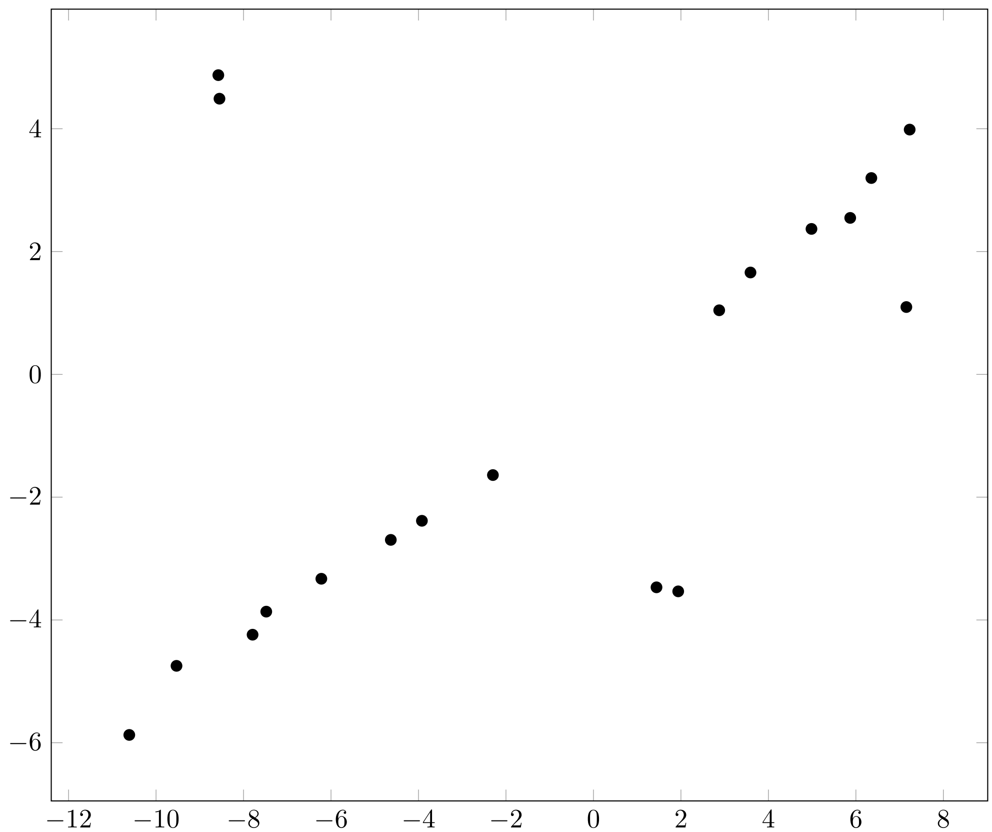
Hierarchical Clustering

There are many variants of hierarchical clustering; here we explore 3. The key difference is how you measure the distance d(S_1, S_2) between two clusters S_1 and S_2.

A (30 points): Run all hierarchical clustering variants on data set C1.txt until there are k = 4 clusters, and report the results as sets. It may be useful to do this pictorially.

Single-Link

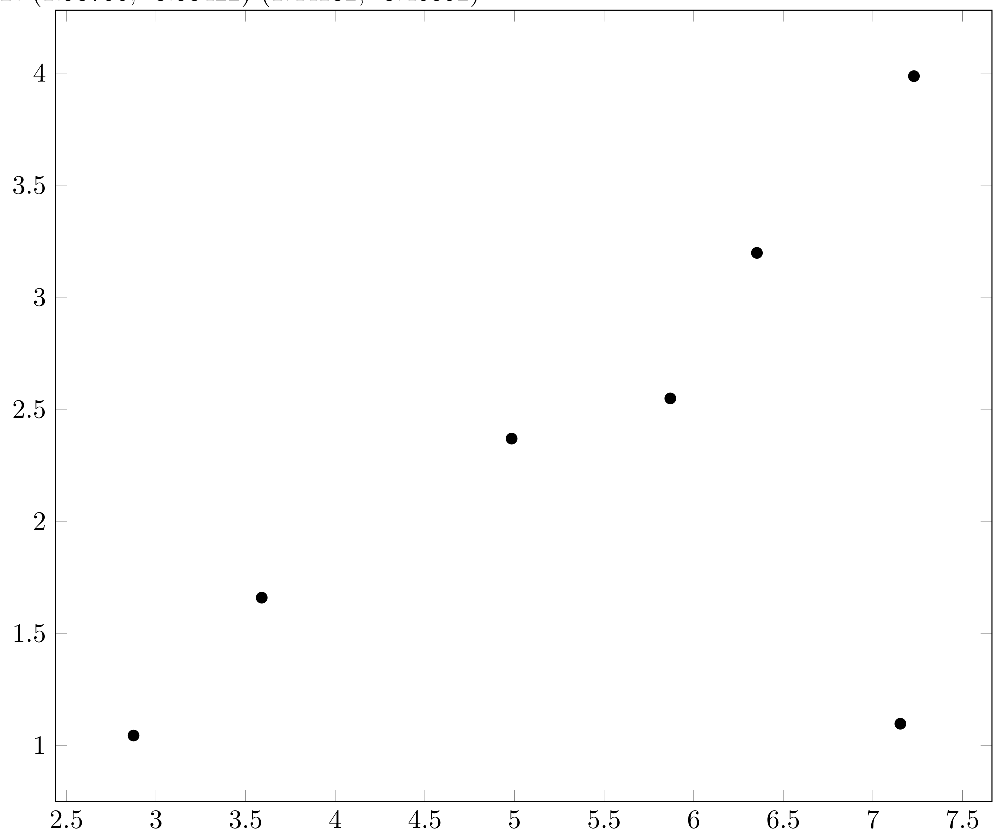


SET: (7.22888, 3.98631) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (3.58997, 1.65892) (2.87479, 1.04368) (7.15324, 1.09632)

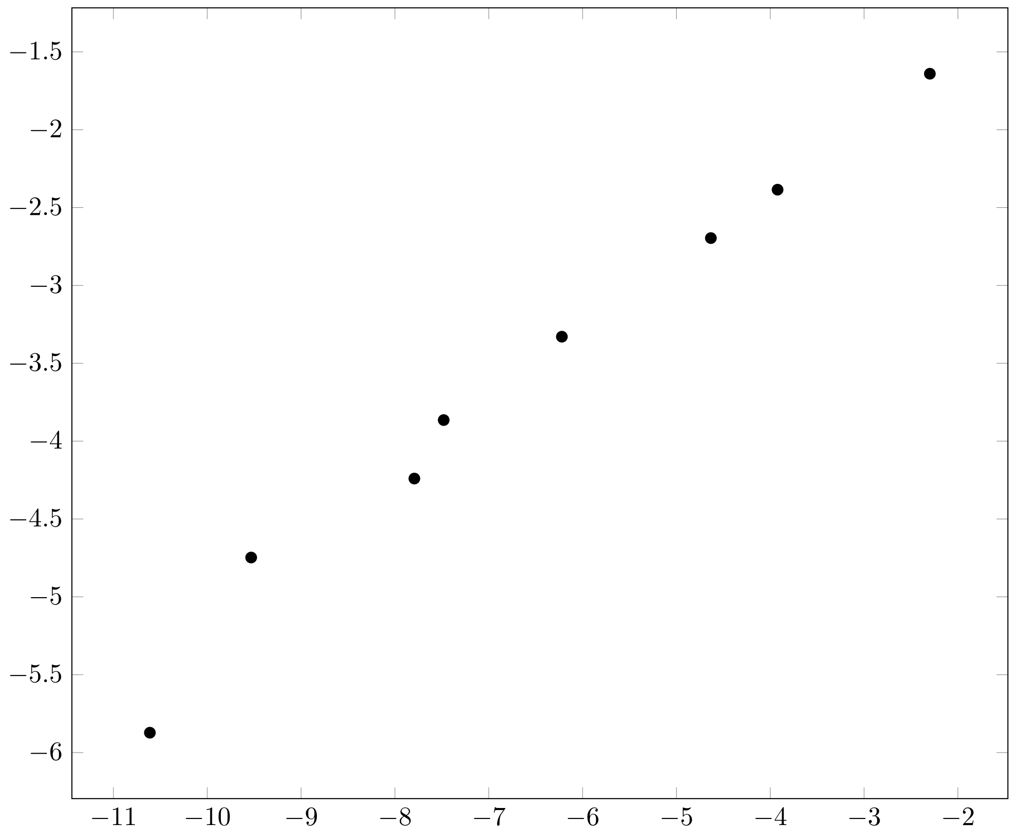
SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-3.92183, -2.38493) (-4.63287, -2.69581) (-2.29879, -1.64012) (-10.61134, -5.87252) (-9.53185, -4.74738)

SET: (-8.57591, 4.87327) (-8.54976, 4.48963)

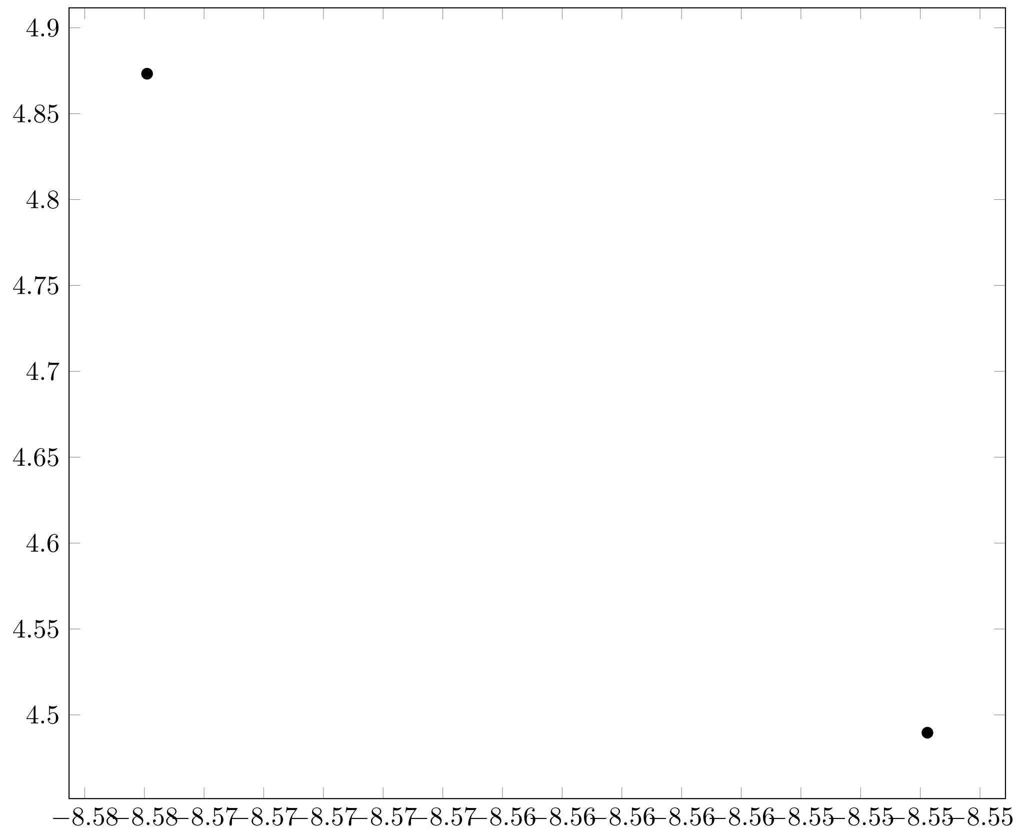
SET: (1.93756, -3.53422) (1.44132, -3.46852)



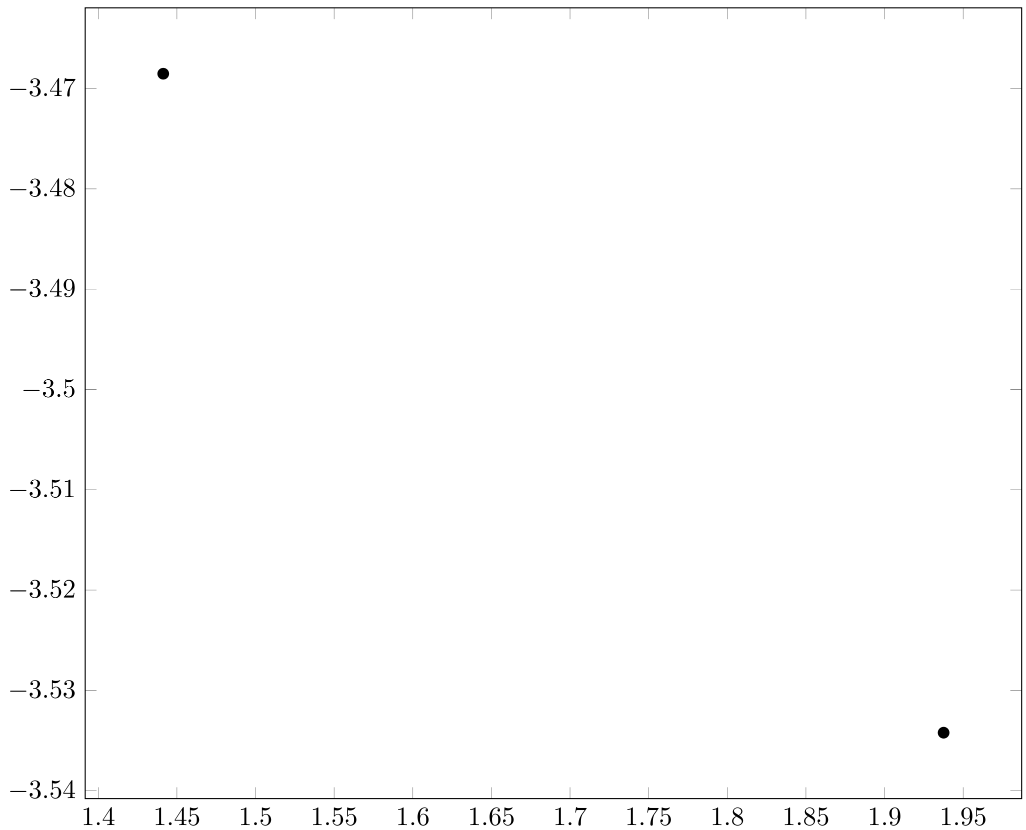
SET: (7.22888, 3.98631) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (3.58997, 1.65892) (2.87479, 1.04368) (7.15324, 1.09632)



SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-3.92183, -2.38493) (-4.63287, -2.69581) (-2.29879, -1.64012) (-10.61134, -5.87252) (-9.53185, -4.74738

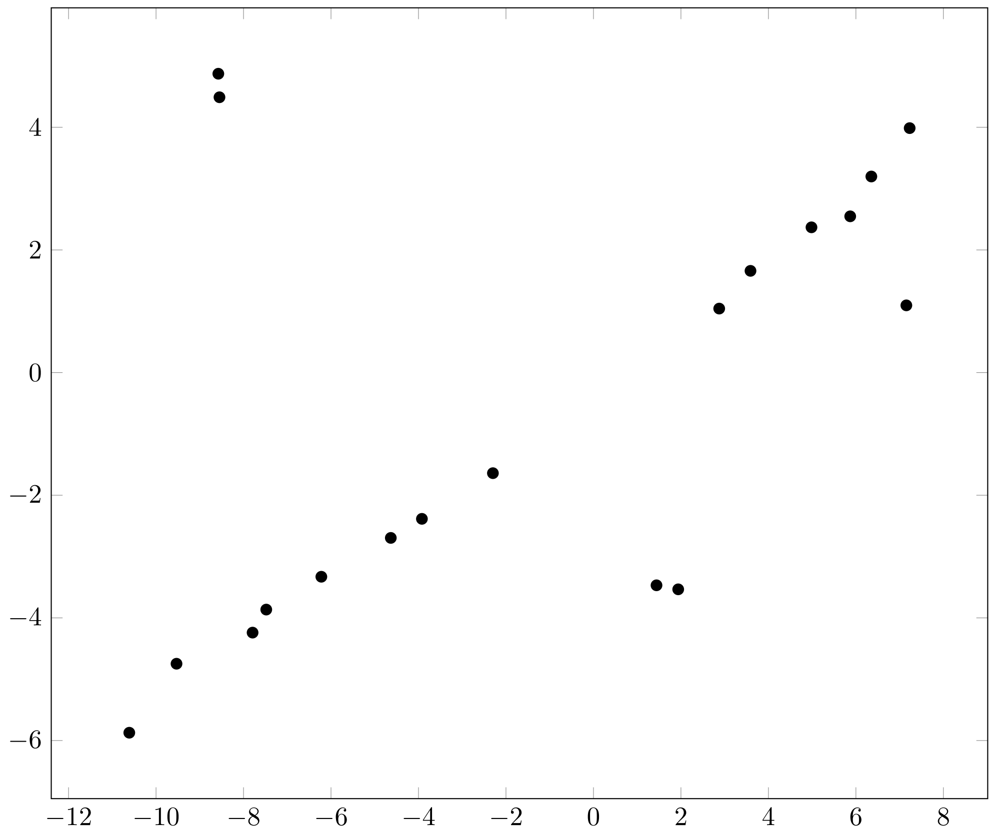


SET: (-8.57591, 4.87327) (-8.54976, 4.48963)



SET: (1.93756, -3.53422) (1.44132, -3.46852)

Complete-Link

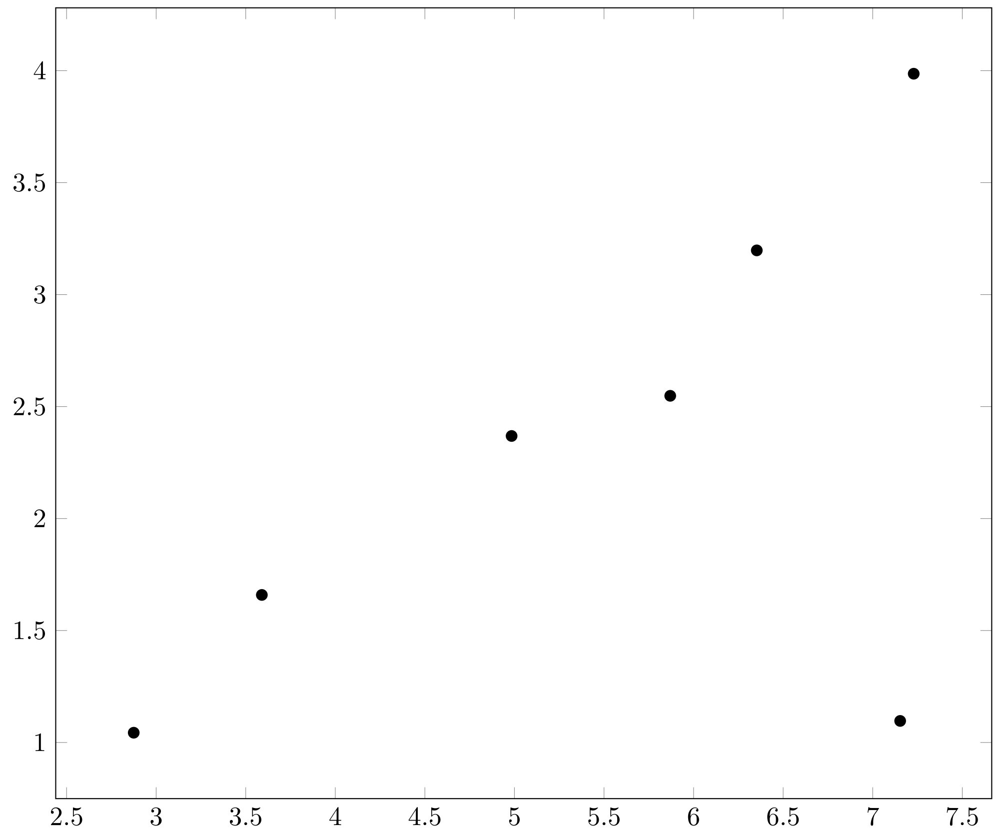


SET: (7.22888, 3.98631) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (7.15324, 1.09632) (3.58997, 1.65892) (2.87479, 1.04368)

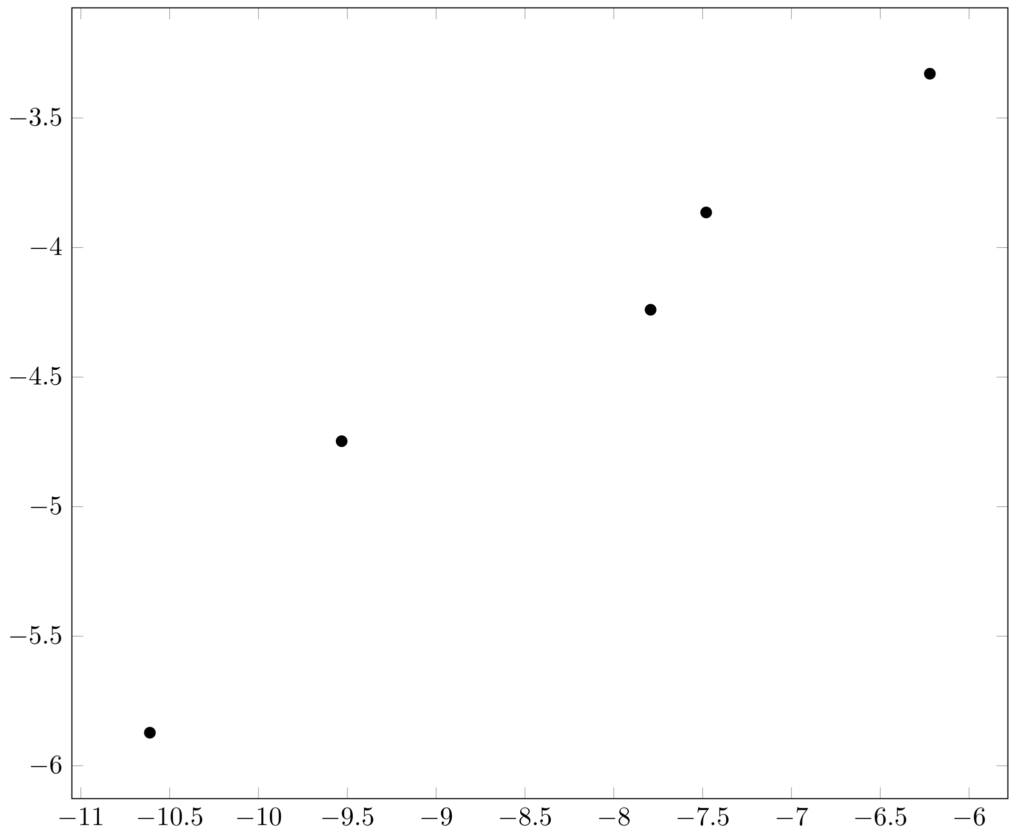
SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-10.61134, -5.87252) (-9.53185, -4.74738)

SET: (-8.57591, 4.87327) (-8.54976, 4.48963)

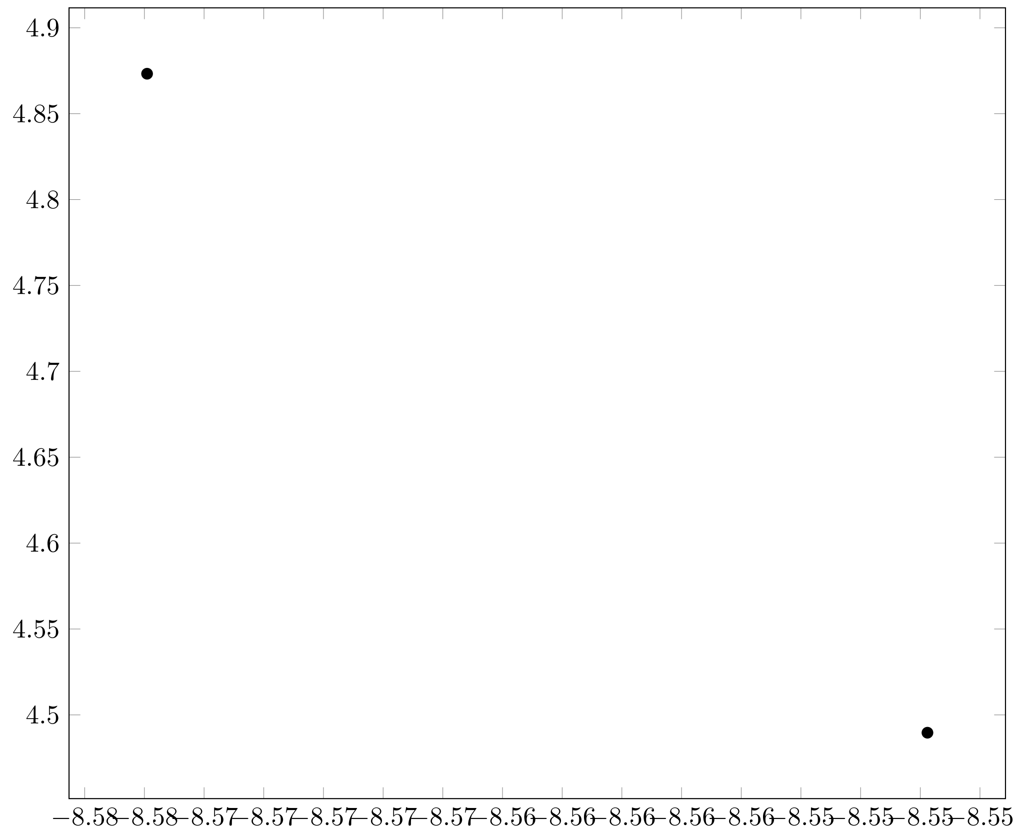
SET: (1.93756, -3.53422) (1.44132, -3.46852) (-3.92183, -2.38493) (-4.63287, -2.69581) (-2.29879, -1.64012)



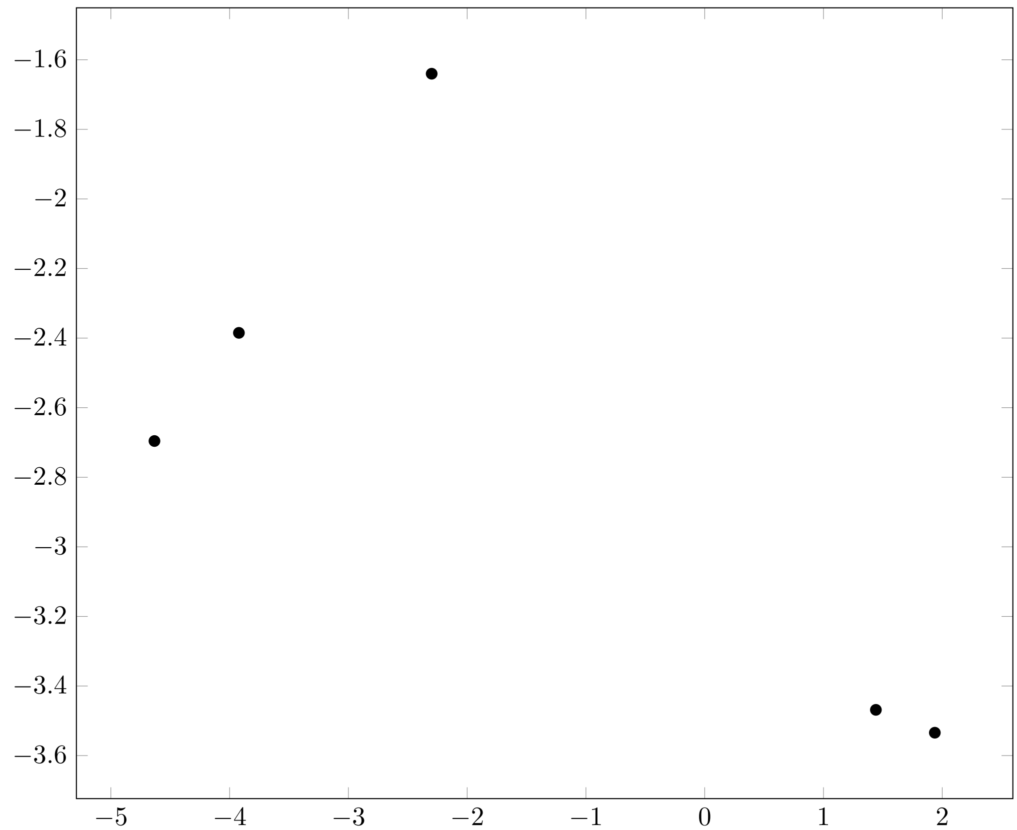
SET: (7.22888, 3.98631) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (7.15324, 1.09632) (3.58997, 1.65892) (2.87479, 1.04368)



SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-10.61134, -5.87252) (-9.53185, -4.74738)

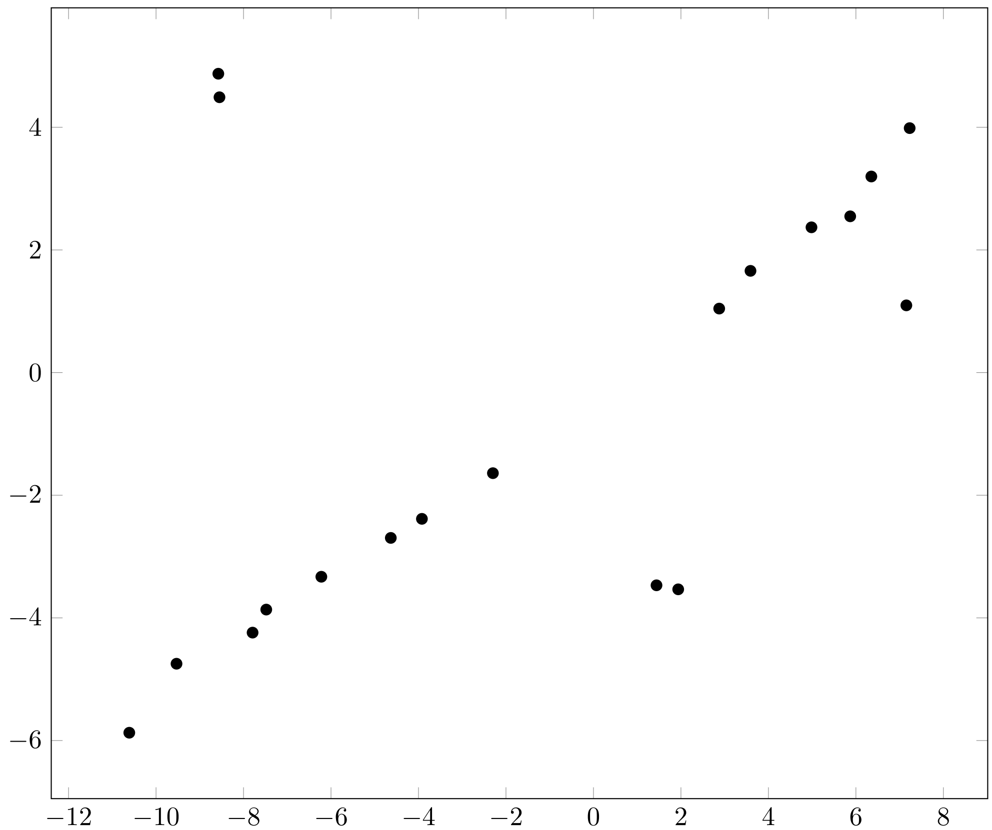


SET: (-8.57591, 4.87327) (-8.54976, 4.48963)



SET: (1.93756, -3.53422) (1.44132, -3.46852) (-3.92183, -2.38493) (-4.63287, -2.69581) (-2.29879, -1.64012)

Mean-Link

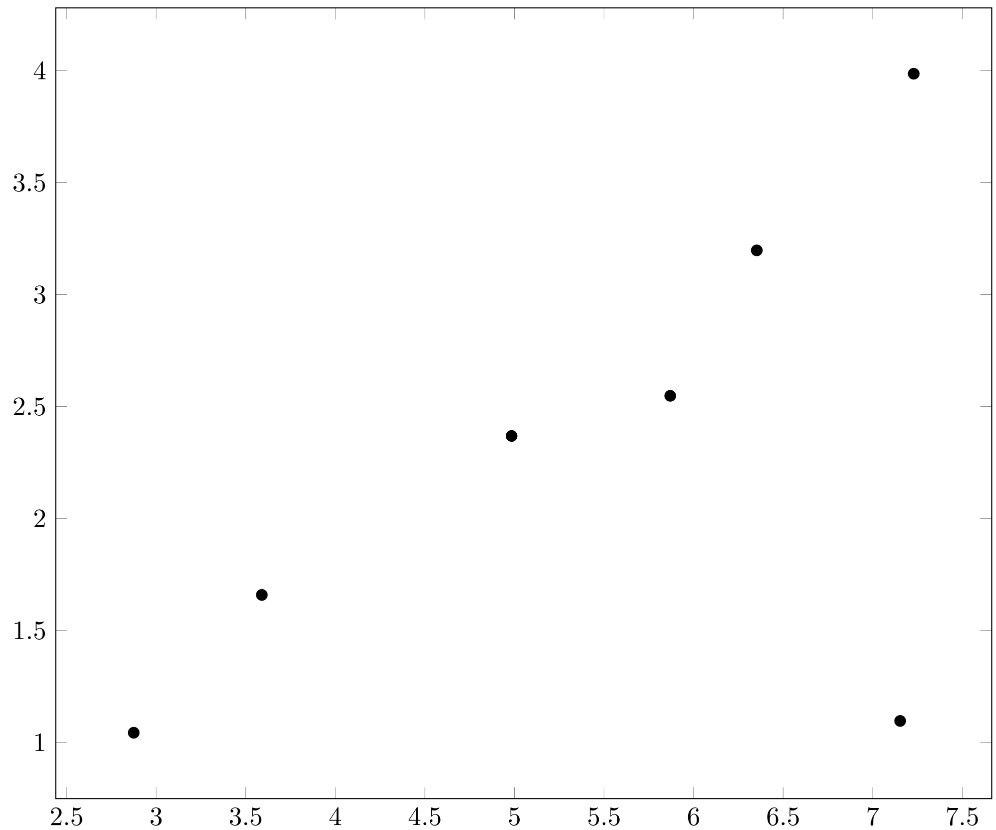


SET: (7.22888, 3.98631) (3.58997, 1.65892) (2.87479, 1.04368) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (7.15324, 1.09632)

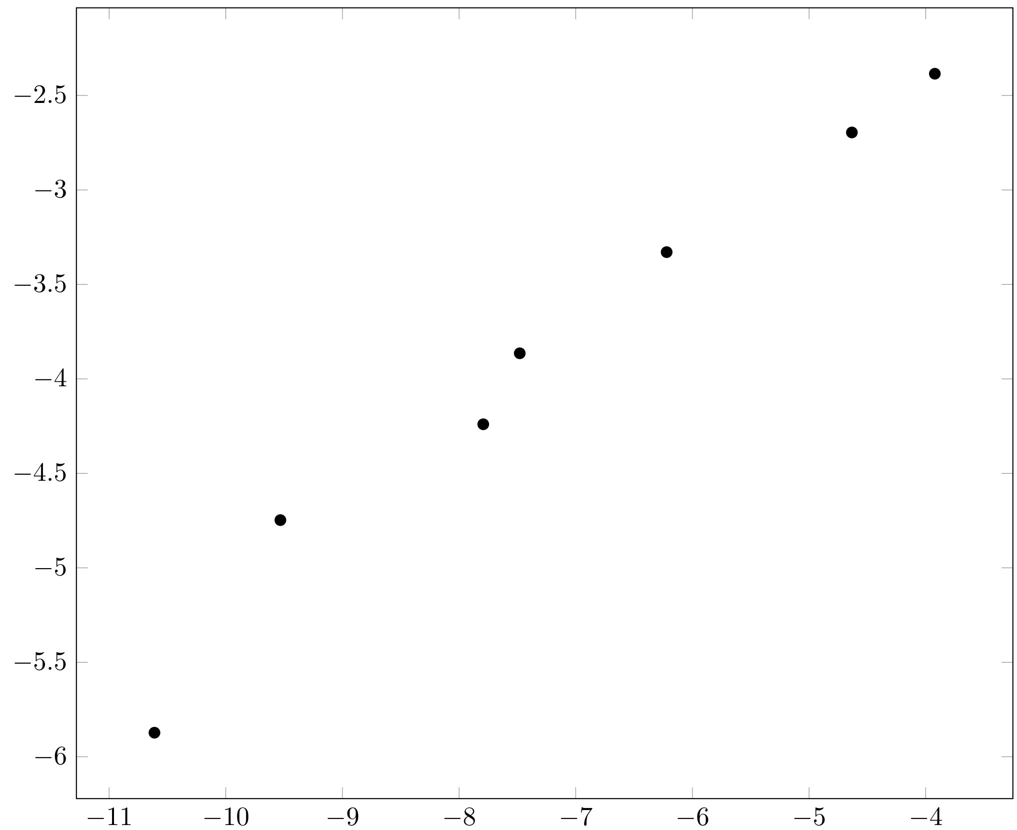
SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-3.92183, -2.38493) (-4.63287, -2.69581) (-9.53185, -4.74738) (-10.61134, -5.87252)

SET: (-8.57591, 4.87327) (-8.54976, 4.48963)

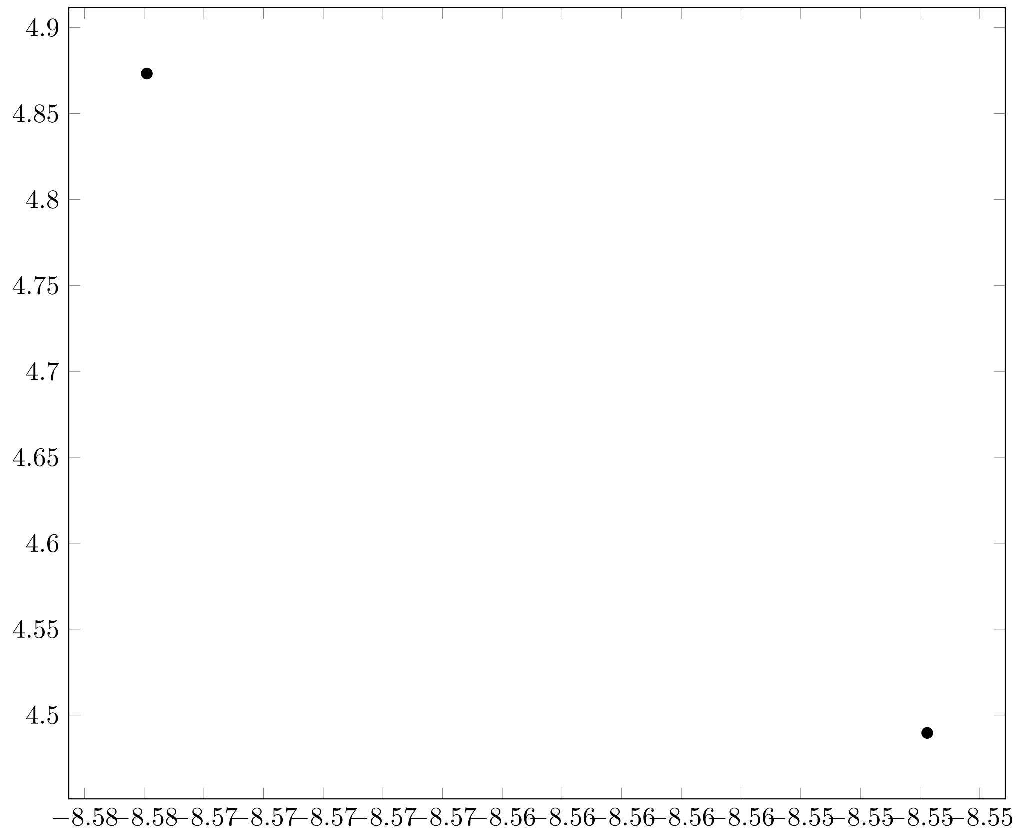
SET: (1.93756, -3.53422) (1.44132, -3.46852) (-2.29879, -1.64012)



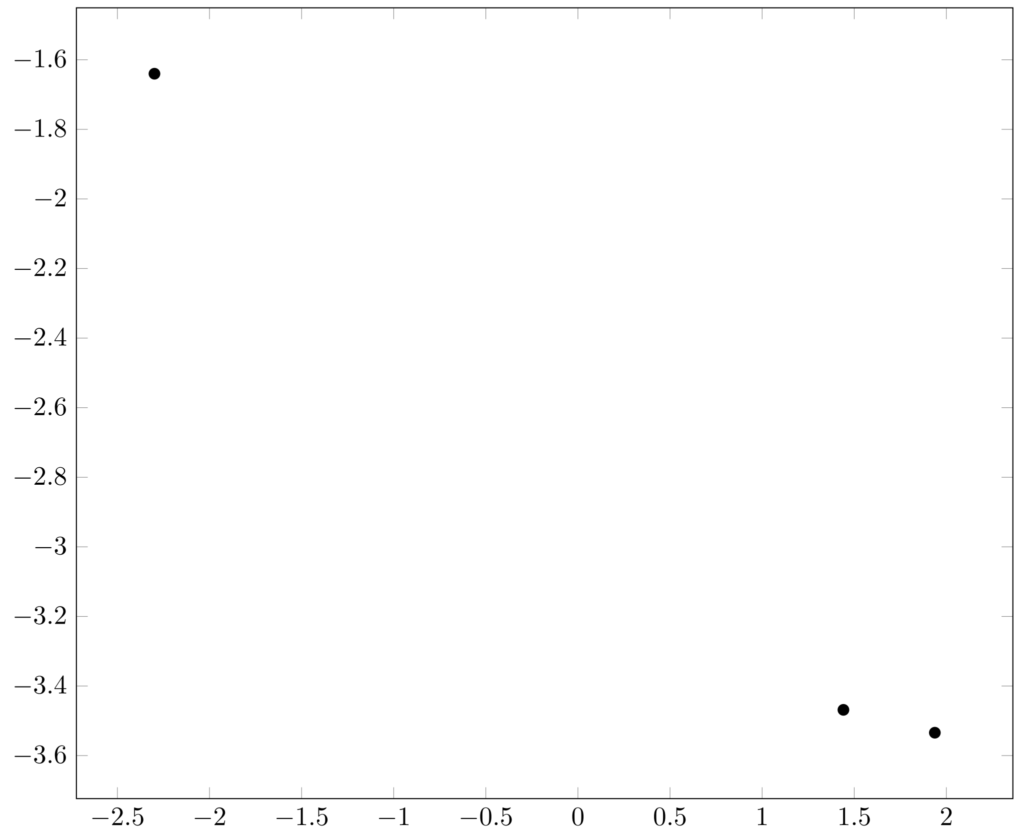
SET: (7.22888, 3.98631) (3.58997, 1.65892) (2.87479, 1.04368) (4.98418, 2.36875) (6.35289, 3.19749) (5.86972, 2.5481) (7.15324, 1.09632)



SET: (-7.79282, -4.24013) (-7.47999, -3.86454) (-6.22075, -3.32916) (-3.92183, -2.38493) (-4.63287, -2.69581) (-9.53185, -4.74738) (-10.61134, -5.87252)



SET: (-8.57591, 4.87327) (-8.54976, 4.48963)



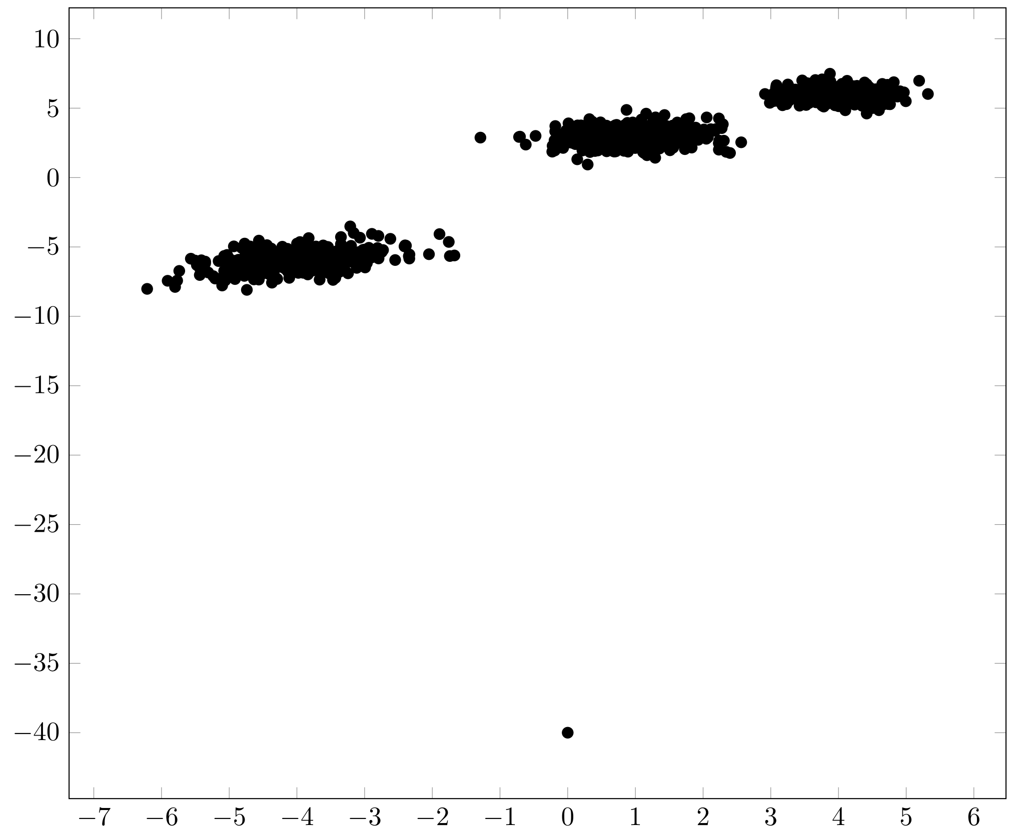
SET: (1.93756, -3.53422) (1.44132, -3.46852) (-2.29879, -1.64012)

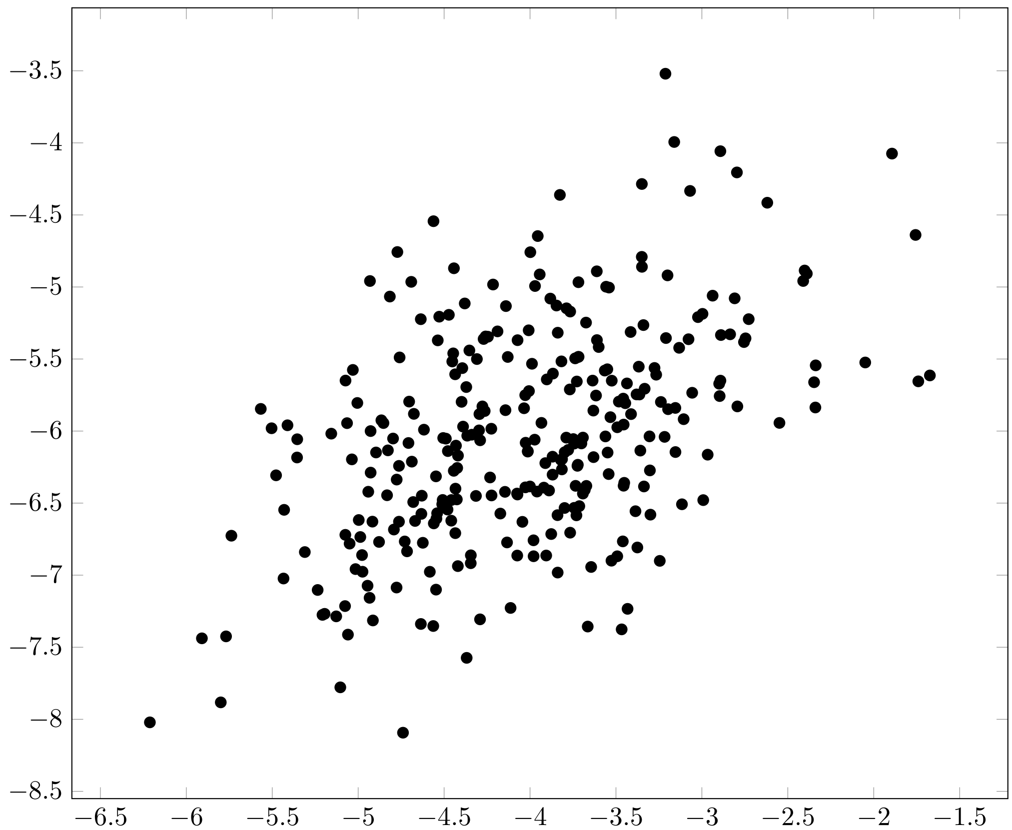
B (5 points): Which variant did the best job, and which was the easiest to compute (think if the data was much larger)? Explain your answers.

The variant that did the best job was the Single-Link, the one who calculates with the closest points in the clusters. The one easiest to compute was the Single-Linked, The others took more calculations when I programed it. But this was just in my Code. Im pretty sure there are more affective ways to code this clustering but mine was just was the Single-Linked one. I Also feel like it was correct enough to be very accurate as well.

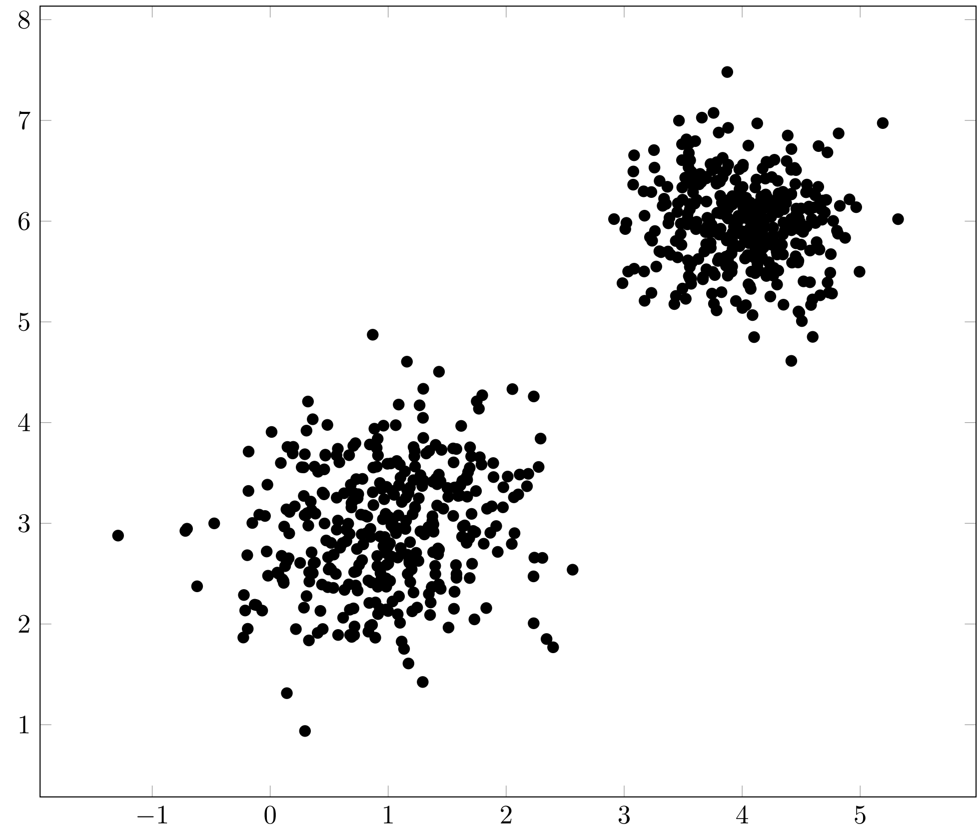
Assignment-Based Clustering

A: (15 points) Run Gonzalez and k-Means++ on data set C2.txt for k = 3. To avoid too much variation in the results, choose c1 as the point with index 1. Report the centers and the subsets (as pictures) for Gonzalez. Report: the 3-center cost and the 3-means cost

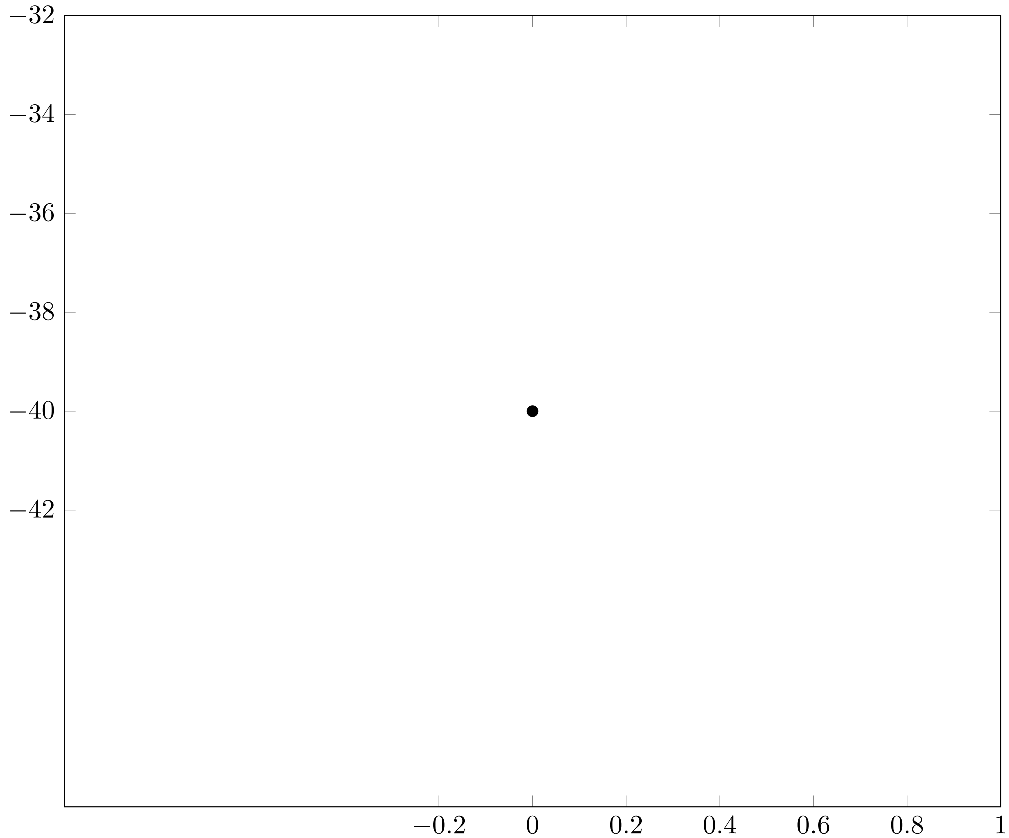

10
5
0
−5
−10
−15
−20
−25
−30
−35
−40
−7 −6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6



Center: (-4.4357274,-5.6060004)

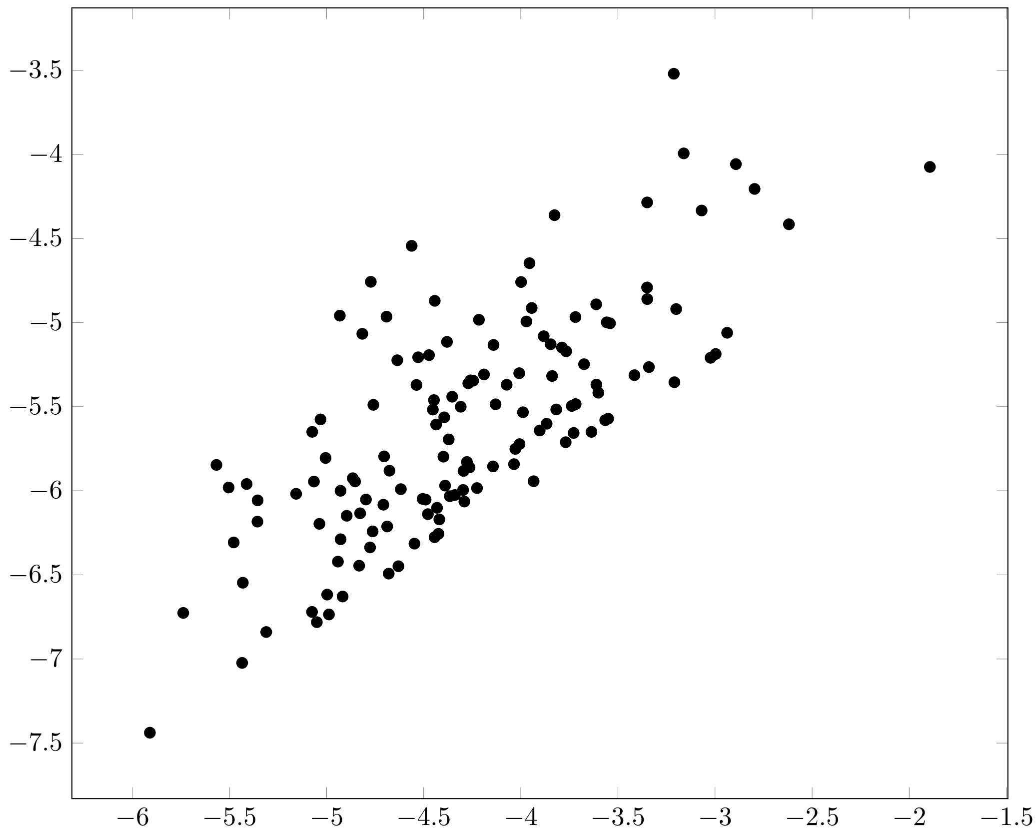


Center: (5.1914717,6.9736254)

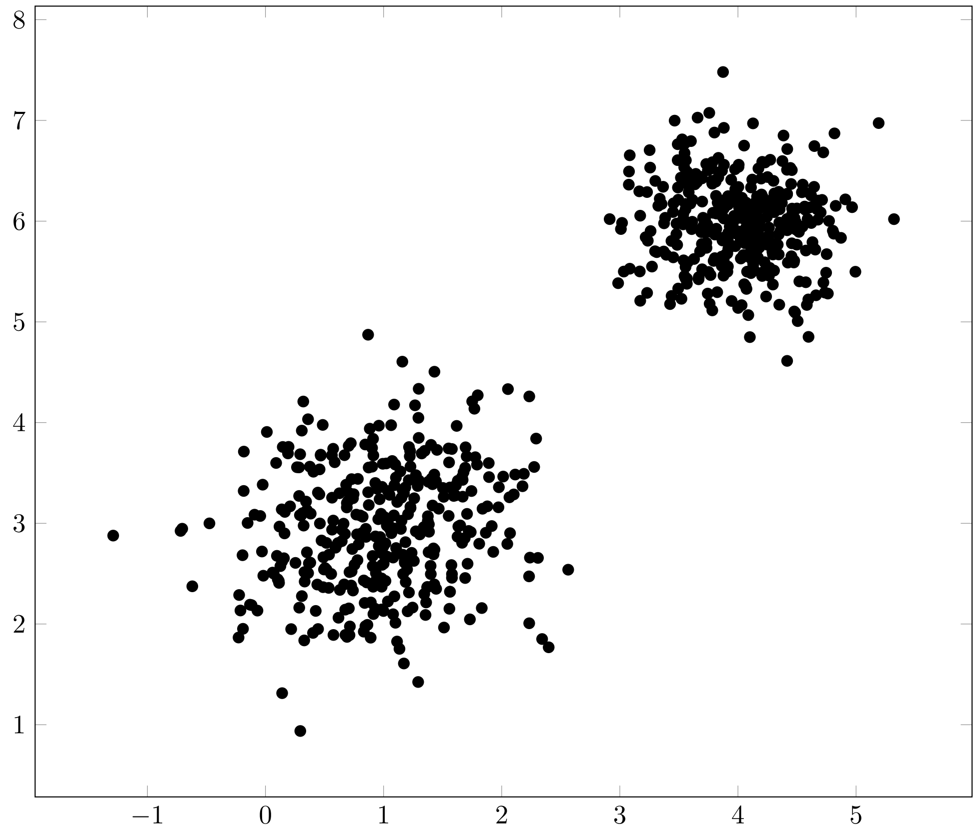


Center: (0.0,-40.0)

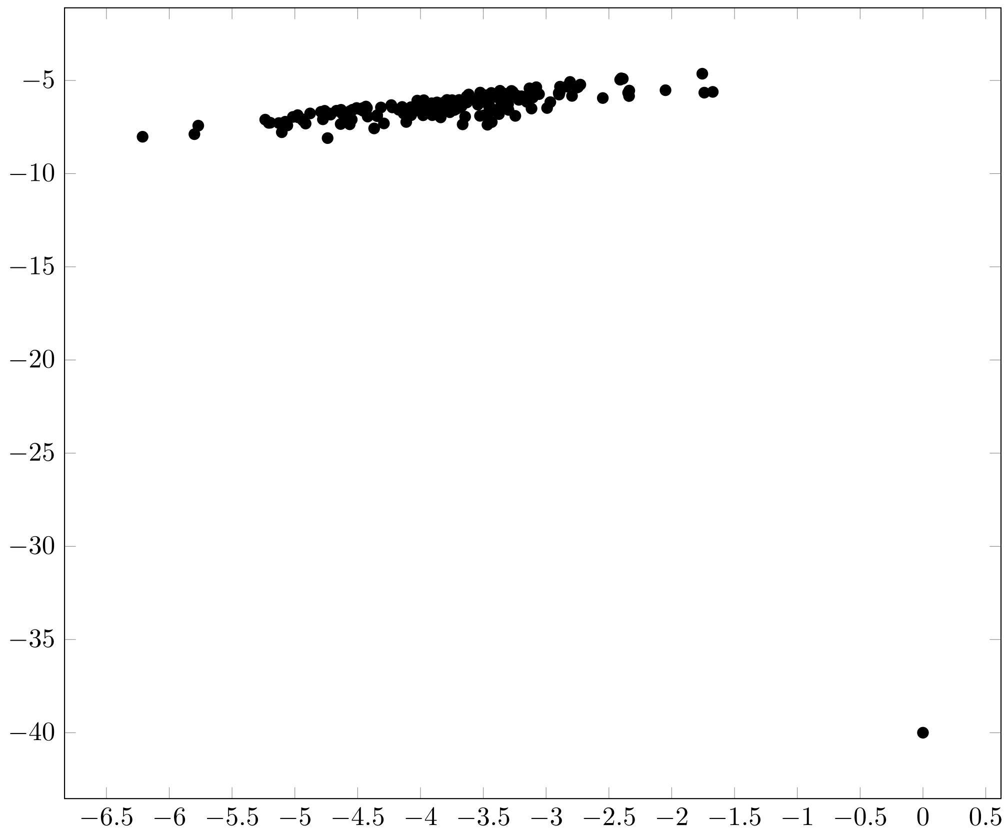
Gonzales
3-center cost = 7.7721092399481115 and 3-means cost = 3.69



Center point at index 1



Center point at index 493

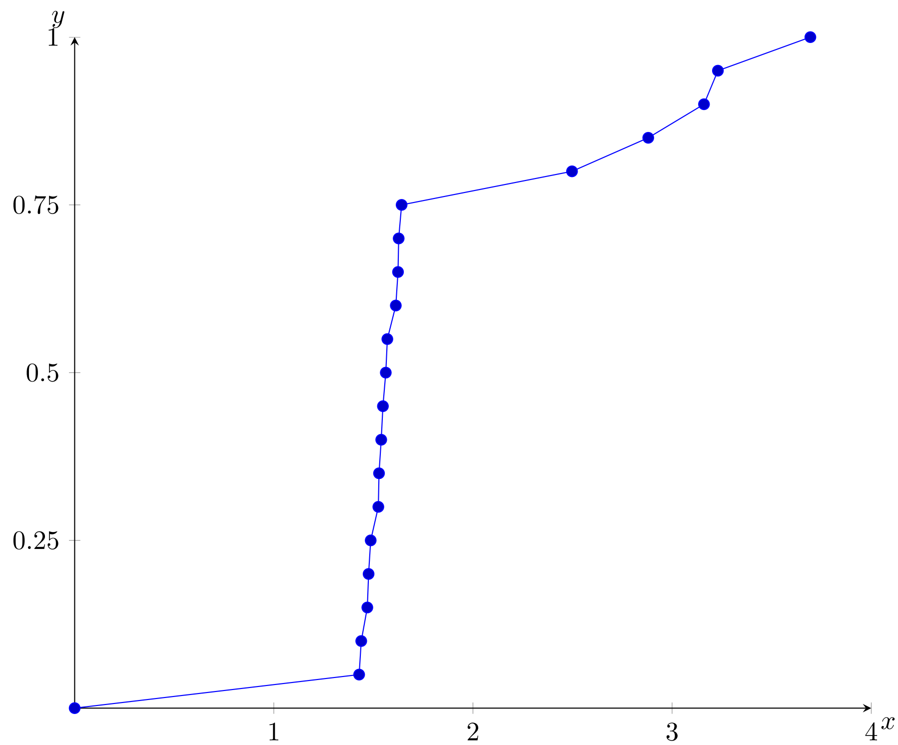


Center point at index 83

k-means++
3-center cost = 34.67885358618291 and 3-means cost = 1.5024098008133218
The result can change sense the result are partially random, this was 1 run but a different run produces different results

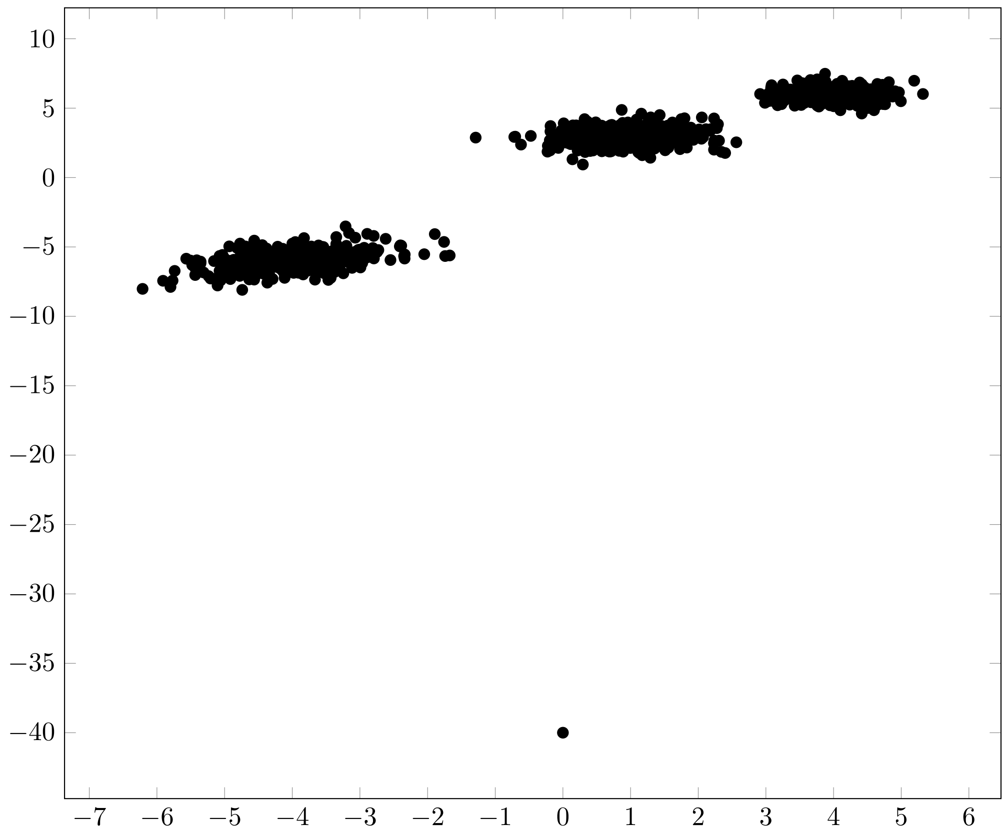
B: (20 points) For k-Means++, the algorithm is randomized, so you will need to report the variation in this algorithm. Run it several trials (at least 20) and plot the cumulative density function of the 3-means cost. Also report what fraction of the time the subsets are the same as the result from Gonzalez.

0/20 times are the set the same, some times they were close, but never the same.

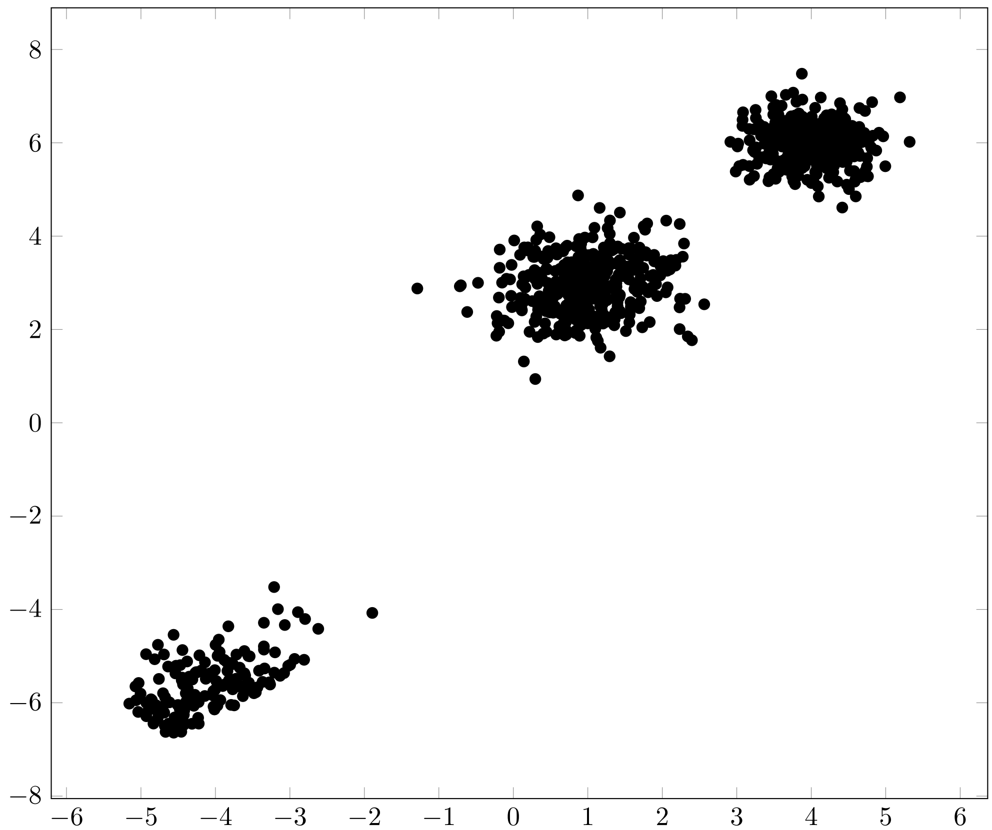


3-means cost = 1.46434

C: (30 points) Recall that Lloyd's algorithm for k-means clustering starts with a set of k centers C and runs as described in Algorithm 9.2.1.
Run Lloyds Algorithm with C initially with points indexed 1,2,3. Report the final subset and the 3-means cost.

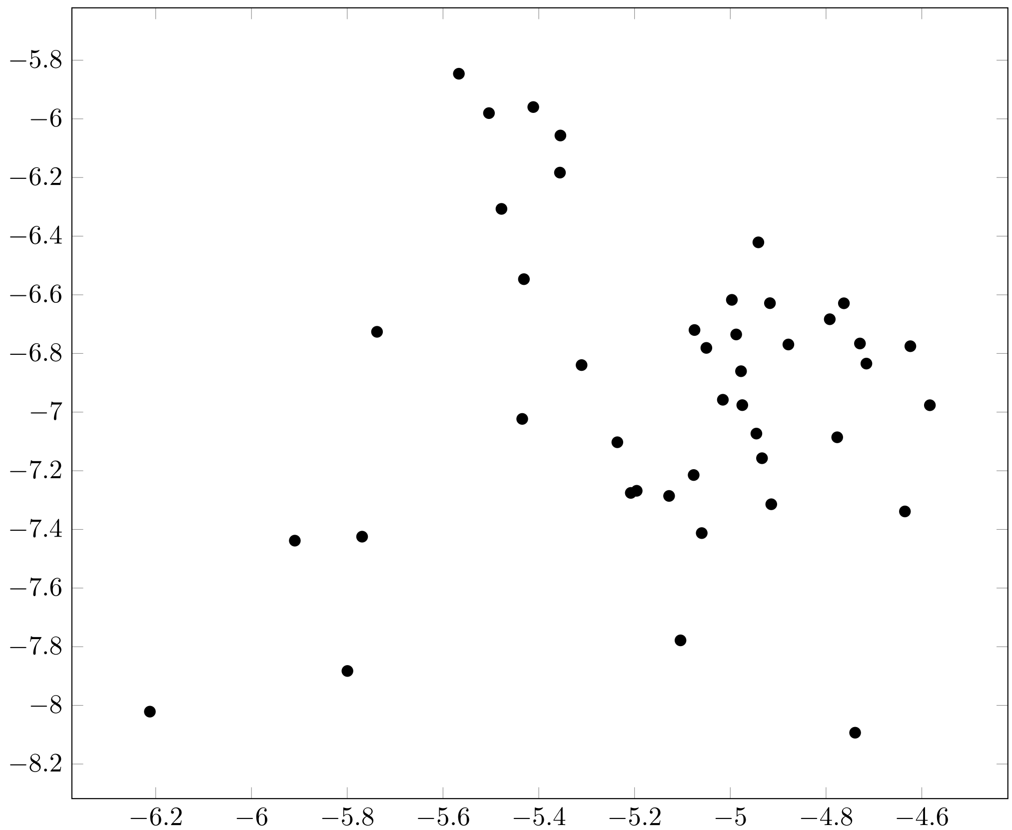


10
5
0
−5
−10
−15
−20
−25
−30
−35
−40
−7
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5
6

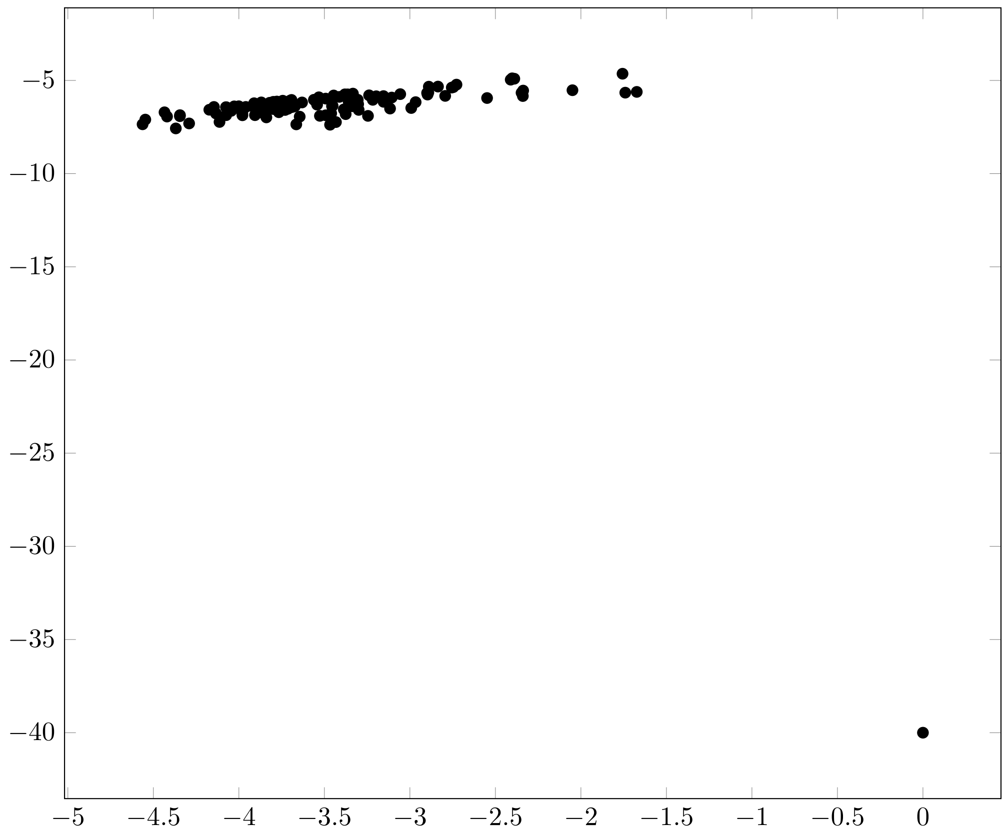


SET:
1 , 4 , 6 , 7 , 8 , 9 , 12 , 14 , 15 , 17 , 18 , 21 , 22 , 23 , 24 , 28 , 30 , 31 , 35 , 40
, 41 , 44 , 46 , 48 , 51 , 53 , 55 , 58 , 59 , 61 , 62 , 67 , 69 , 73 , 74 , 76 , 83 , 87
, 89 , 91 , 92 , 93 , 94 , 95 , 97 , 98 , 99 , 101 , 104 , 105 , 106 , 109 , 110 , 111 ,
115 , 116 , 117 , 120 , 121 , 122 , 124 , 125 , 129 , 130 , 132 , 133 , 135 , 138 ,
139 , 145 , 146 , 150 , 151 , 152 , 153 , 155 , 156 , 157 , 160 , 166 , 167 , 168 ,
170 , 172 , 173 , 176 , 177 , 178 , 181 , 182 , 188 , 192 , 193 , 197 , 198 , 199 ,
202 , 203 , 204 , 205 , 206 , 207 , 208 , 213 , 214 , 215 , 217 , 218 , 219 , 220 ,
222 , 225 , 228 , 229 , 230 , 231 , 233 , 234 , 235 , 237 , 239 , 241 , 245 , 247 ,
248 , 249 , 251 , 252 , 254 , 261 , 262 , 263 , 265 , 271 , 272 , 273 , 274 , 275 ,
276 , 277 , 278 , 279 , 281 , 282 , 285 , 286 , 289 , 290 , 293 , 296 , 298 , 299 ,
302 , 305 , 306 , 307 , 308 , 309 , 310 , 311 , 312 , 313 , 314 , 315 , 316 , 317 ,
318 , 319 , 320 , 321 , 322 , 323 , 324 , 325 , 326 , 327 , 328 , 329 , 330 , 331 ,
332 , 333 , 334 , 335 , 336 , 337 , 338 , 339 , 340 , 341 , 342 , 343 , 344 , 345 ,
346 , 347 , 348 , 349 , 350 , 351 , 352 , 353 , 354 , 355 , 356 , 357 , 358 , 359 ,
360 , 361 , 362 , 363 , 364 , 365 , 366 , 367 , 368 , 369 , 370 , 371 , 372 , 373 ,
374 , 375 , 376 , 377 , 378 , 379 , 380 , 381 , 382 , 383 , 384 , 385 , 386 , 387 ,
388 , 389 , 390 , 391 , 392 , 393 , 394 , 395 , 396 , 397 , 398 , 399 , 400 , 401 ,
402 , 403 , 404 , 405 , 406 , 407 , 408 , 409 , 410 , 411 , 412 , 413 , 414 , 415 ,

416 , 417 , 418 , 419 , 420 , 421 , 422 , 423 , 424 , 425 , 426 , 427 , 428 , 429 ,
430 , 431 , 432 , 433 , 434 , 435 , 436 , 437 , 438 , 439 , 440 , 441 , 442 , 443 ,
444 , 445 , 446 , 447 , 448 , 449 , 450 , 451 , 452 , 453 , 454 , 455 , 456 , 457 ,
458 , 459 , 460 , 461 , 462 , 463 , 464 , 465 , 466 , 467 , 468 , 469 , 470 , 471 ,
472 , 473 , 474 , 475 , 476 , 477 , 478 , 479 , 480 , 481 , 482 , 483 , 484 , 485 ,
486 , 487 , 488 , 489 , 490 , 491 , 492 , 493 , 494 , 495 , 496 , 497 , 498 , 499 ,
500 , 501 , 502 , 503 , 504 , 505 , 506 , 507 , 508 , 509 , 510 , 511 , 512 , 513 ,
514 , 515 , 516 , 517 , 518 , 519 , 520 , 521 , 522 , 523 , 524 , 525 , 526 , 527 ,
528 , 529 , 530 , 531 , 532 , 533 , 534 , 535 , 536 , 537 , 538 , 539 , 540 , 541 ,
542 , 543 , 544 , 545 , 546 , 547 , 548 , 549 , 550 , 551 , 552 , 553 , 554 , 555 ,
556 , 557 , 558 , 559 , 560 , 561 , 562 , 563 , 564 , 565 , 566 , 567 , 568 , 569 ,
570 , 571 , 572 , 573 , 574 , 575 , 576 , 577 , 578 , 579 , 580 , 581 , 582 , 583 ,
584 , 585 , 586 , 587 , 588 , 589 , 590 , 591 , 592 , 593 , 594 , 595 , 596 , 597 ,
598 , 599 , 600 , 601 , 602 , 603 , 604 , 605 , 606 , 607 , 608 , 609 , 610 , 611 ,
612 , 613 , 614 , 615 , 616 , 617 , 618 , 619 , 620 , 621 , 622 , 623 , 624 , 625 ,
626 , 627 , 628 , 629 , 630 , 631 , 632 , 633 , 634 , 635 , 636 , 637 , 638 , 639 ,
640 , 641 , 642 , 643 , 644 , 645 , 646 , 647 , 648 , 649 , 650 , 651 , 652 , 653 ,
654 , 655 , 656 , 657 , 658 , 659 , 660 , 661 , 662 , 663 , 664 , 665 , 666 , 667 ,
668 , 669 , 670 , 671 , 672 , 673 , 674 , 675 , 676 , 677 , 678 , 679 , 680 , 681 ,
682 , 683 , 684 , 685 , 686 , 687 , 688 , 689 , 690 , 691 , 692 , 693 , 694 , 695 ,
696 , 697 , 698 , 699 , 700 , 701 , 702 , 703 , 704 , 705 , 706 , 707 , 708 , 709 ,
710 , 711 , 712 , 713 , 714 , 715 , 716 , 717 , 718 , 719 , 720 , 721 , 722 , 723 ,
724 , 725 , 726 , 727 , 728 , 729 , 730 , 731 , 732 , 733 , 734 , 735 , 736 , 737 ,
738 , 739 , 740 , 741 , 742 , 743 , 744 , 745 , 746 , 747 , 748 , 749 , 750 , 751 ,
752 , 753 , 754 , 755 , 756 , 757 , 758 , 759 , 760 , 761 , 762 , 763 , 764 , 765 ,
766 , 767 , 768 , 769 , 770 , 771 , 772 , 773 , 774 , 775 , 776 , 777 , 778 , 779 ,
780 , 781 , 782 , 783 , 784 , 785 , 786 , 787 , 788 , 789 , 790 , 791 , 792 , 793 ,
794 , 795 , 796 , 797 , 798 , 799 , 800 , 801 , 802 , 803 , 804 , 805 , 806 , 807 ,
808 , 809 , 810 , 811 , 812 , 813 , 814 , 815 , 816 , 817 , 818 , 819 , 820 , 821 ,
822 , 823 , 824 , 825 , 826 , 827 , 828 , 829 , 830 , 831 , 832 , 833 , 834 , 835 ,
836 , 837 , 838 , 839 , 840 , 841 , 842 , 843 , 844 , 845 , 846 , 847 , 848 , 849 ,
850 , 851 , 852 , 853 , 854 , 855 , 856 , 857 , 858 , 859 , 860 , 861 , 862 , 863 ,
864 , 865 , 866 , 867 , 868 , 869 , 870 , 871 , 872 , 873 , 874 , 875 , 876 , 877 ,
878 , 879 , 880 , 881 , 882 , 883 , 884 , 885 , 886 , 887 , 888 , 889 , 890 , 891 ,
892 , 893 , 894 , 895 , 896 , 897 , 898 , 899 , 900 , 901 , 902 , 903 , 904 , 905 ,
906 , 907 , 908 , 909 , 910 , 911 , 912 , 913 , 914 , 915 , 916 , 917 , 918 , 919 ,
920 , 921 , 922 , 923 , 924 , 925 , 926 , 927 , 928 , 929 , 930 , 931 , 932 , 933 ,
934 , 935 , 936 , 937 , 938 , 939 , 940 , 941 , 942 , 943 , 944 , 945 , 946 , 947 ,
948 , 949 , 950 , 951 , 952 , 953 , 954 , 955 , 956 , 957 , 958 , 959 , 960 , 961 ,
962 , 963 , 964 , 965 , 966 , 967 , 968 , 969 , 970 , 971 , 972 , 973 , 974 , 975 ,
976 , 977 , 978 , 979 , 980 , 981 , 982 , 983 , 984 , 985 , 986 , 987 , 988 , 989 ,
990 , 991 , 992 , 993 , 994 , 995 , 996 , 997 , 998 , 999 , 1000 , 1001 , 1002 ,
1003 , 1004 , 1005 , 1006 , 1007 , 1008 , 1009 , 1010 , 1011 , 1012 , 1013 , 1014
, 1015 , 1016 , 1017 , 1018 , 1019 , 1020 , 1021 , 1022 , 1023 , 1024 , 1025 ,
1026 , 1027 , 1028 , 1029 , 1030 , 1031 , 1032 , 1033 , 1034 , 1035 , 1036 , 1037
, 1038 , 1039 ,



SET:
(-5.7378081,-6.7261898)(-5.5040044,-5.9804901)(-5.1954559,-7.2683139)(-4.6239007,-6.7752677)(-5.1039726,-7.7783371)(-5.9094735,-7.4383142)(-5.2077795,-7.2755335)(-5.5668423,-5.8461305)(-4.9876562,-6.7351985)(-5.4310054,-6.5466759)(-5.0596488,-7.412555)(-4.7627258,-6.6288991)(-4.7157247,-6.8345469)(-5.3557532,-6.1833345)(-5.4345237,-7.0232062)(-4.9337278,-7.1572124)(-4.9749228,-6.9761827)(-5.0747144,-6.7200857)(-5.2358732,-7.1027816)(-5.1278103,-7.2858295)(-4.776674,-7.0858312)(-6.2121825,-8.0213802)(-4.9413387,-6.4210905)(-5.0499548,-6.7810673)(-5.7996125,-7.8827268)(-4.7922369,-6.6833058)(-4.9171777,-6.6284922)(-4.6352562,-7.3386924)(-4.5831507,-6.9765101)(-4.9967003,-6.6172267)(-5.0764888,-7.2145599)(-4.7291797,-6.7660214)(-5.4779079,-6.3070256)(-4.8786631,-6.769412)(-4.9775755,-6.8605382)(-5.7689619,-7.4247345)(-5.0155156,-6.9578993)(-4.9454794,-7.07323)(-5.4114304,-5.9597791)(-4.7392768,-8.0932025)(-4.9144064,-7.3142177)(-5.354781,-6.056786)(-5.31057,-6.8396423)

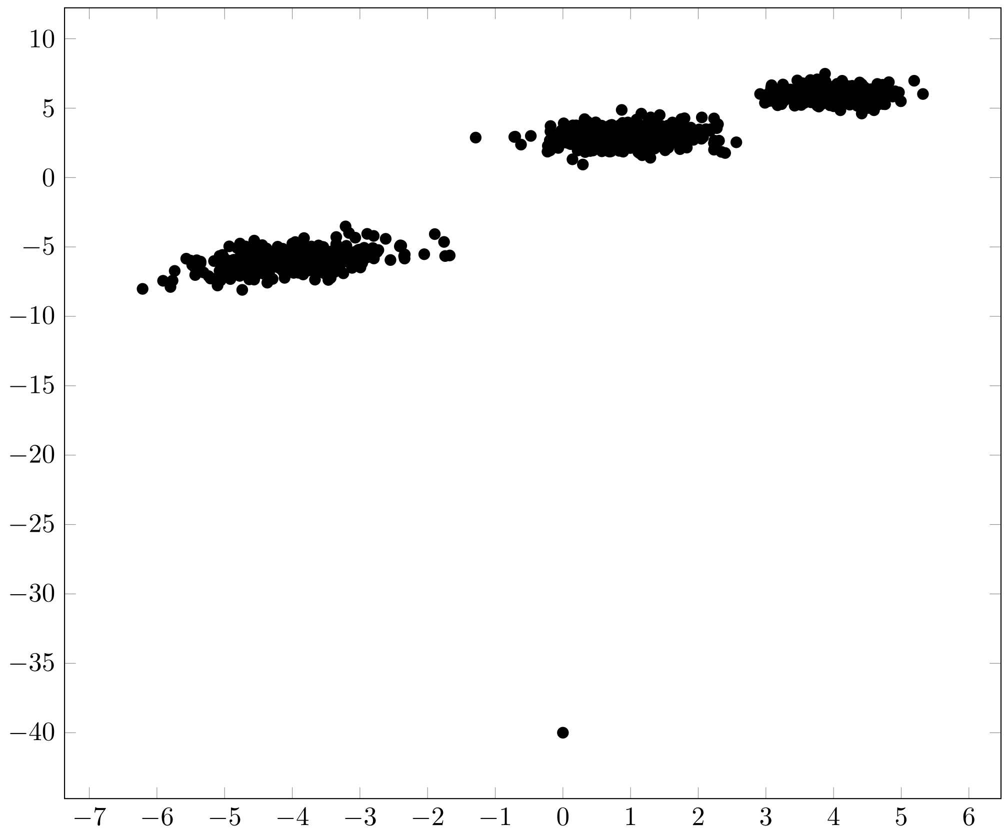


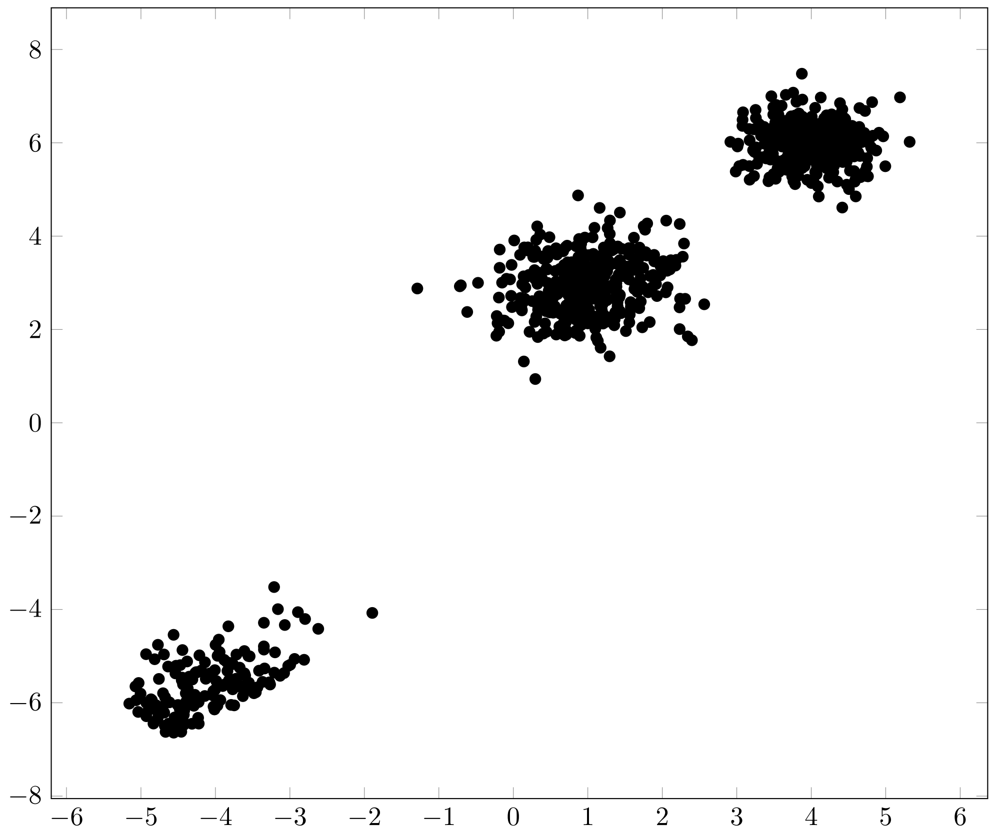
SET:
(-3.3755073,-6.8078161)(-3.7241204,-6.2411353)(-3.1538396,-6.1455886)(-
4.0004525,-6.384472)(-2.8994207,-5.6713194)(-3.7348412,-6.3800545)(-
3.4456156,-5.8066268)(-3.7434374,-6.0869206)(-3.8148002,-6.2660256)(-
1.756621,-4.6388883)(-3.8688091,-6.302476)(-3.8893691,-6.4123887)(-3.692487,-
6.4337346)(-3.8689858,-6.1779057)(-2.3395741,-5.8363895)(-4.2909707,-
7.3065739)(-2.7551086,-5.3823124)(-3.5262238,-6.8995217)(-3.6305771,-
6.1808633)(-3.7300596,-6.5847434)(-4.3687402,-7.5738585)(-3.7780393,-
6.1308289)(-3.1543975,-5.8393938)(-2.3377491,-5.544014)(-3.7429193,-
6.5328615)(-3.304779,-6.0367647)(-3.4924452,-6.8689173)(-3.4552715,-
5.9550063)(-2.7275471,-5.2233346)(-3.3371386,-6.3843846)(-3.7011366,-
6.0857448)(-3.7132858,-6.5209782)(-2.5486799,-5.9433799)(-2.8893272,-
5.3335974)(-4.1458744,-6.421741)(-1.7405971,-5.6548925)(-2.8357246,-
5.3285584)(-3.3802834,-5.7456183)(-3.3330688,-5.7062118)(-4.0444743,-
6.6299313)(-3.8401947,-6.5835544)(-4.5470907,-7.1003811)(-4.1338825,-
6.7728377)(-4.0755467,-6.4335519)(-2.4020712,-4.8858217)(-3.4324807,-
7.2337209)(-3.7219151,-6.2338533)(-2.7460234,-5.3564866)(-4.3445077,-
6.8621011)(-3.7991001,-6.5338046)(-4.4200645,-6.9371697)(-3.79842,-
6.1468452)(-2.3894367,-4.9073916)(-3.6921877,-6.0451664)(-2.7935863,-

5.8288476)(-3.3581338,-6.1354949)(-3.9110232,-6.2223852)(-4.0272515,-6.3906384)(-2.9663929,-6.1638234)(-4.4348705,-6.7084391)(-3.9058857,-6.8635963)(-3.5487112,-6.1490269)(-2.9921091,-6.4791852)(-3.9797656,-6.7580299)(-1.673313,-5.6139879)(-2.892201,-5.6500006)(-4.1132262,-7.2274724)(-3.8392174,-6.9819397)(-3.1160021,-6.5080383)(-4.0744583,-6.8638336)(-3.4668627,-7.3761174)(-3.1967797,-5.8481289)(-3.5613812,-6.0371348)(-4.5631165,-7.3528581)(-4.1728637,-6.5724724)(-2.3471119,-5.6609861)(-3.3025399,-6.2729677)(-3.8137563,-6.194642)(-2.4105049,-4.9584442)(-3.672274,-6.3804859)(-3.4923589,-5.9735848)(-3.4555256,-6.3792907)(-3.2994779,-6.5800543)(-3.6813639,-6.4080051)(-3.4518845,-6.3591688)(-4.0731858,-6.4402769)(-3.3864362,-6.5561329)(-3.8204138,-6.2000687)(-3.5425328,-6.2983828)(-3.2450639,-6.9007445)(-3.2386227,-5.798026)(-3.5322929,-5.9035543)(-3.2169952,-6.0402687)(-4.3450402,-6.9170209)(-2.0494089,-5.5240076)(-3.6643534,-7.3565747)(-2.8966174,-5.757131)(-3.6443604,-6.9431494)(-3.919459,-6.3924113)(-3.3634724,-5.7461759)(-3.1061304,-5.9175077)(-3.7665645,-6.7051905)(-3.9792969,-6.8686698)(-3.876685,-6.7145541)(-3.4601547,-6.7655543)(-3.0564455,-5.7342068)(-3.4123226,-5.88231)(-3.959547,-6.419806)(0.0,-40.0)

3-means cost = 2.216
centers: (2.49628862 4.47085919), (-4.63876365 -6.94026971),(-3.60701252 -5.62649969)

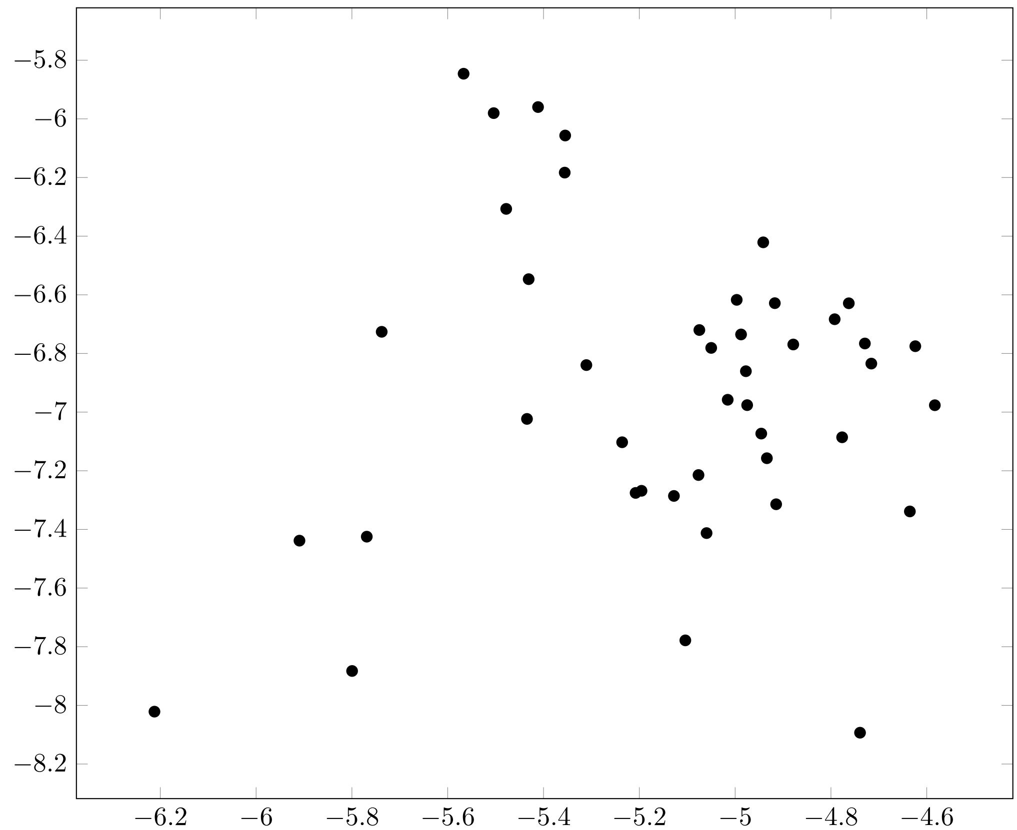
Run Lloyds Algorithm with C initially as the output of Gonzalez above. Report the final subset and the 3-means cost.


10
5
0
−5
−10
−15
−20
−25
−30
−35
−40
−7
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5
6

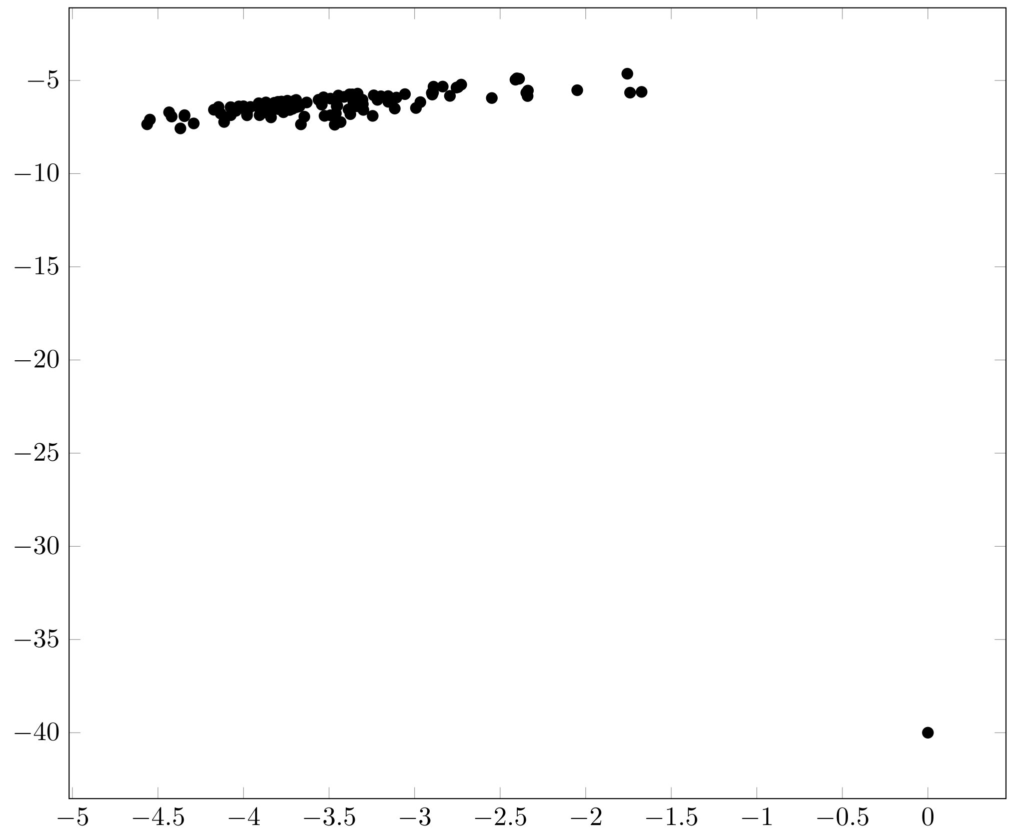


SET:
1 , 4 , 6 , 7 , 8 , 9 , 12 , 14 , 15 , 17 , 18 , 21 , 22 , 23 , 24 , 28 , 30 , 31 , 35 , 40 , 41 , 44 , 46 , 48 , 51 , 53 , 55 , 58 , 59 , 61 , 62 , 67 , 69 , 73 , 74 , 76 , 83 , 87 , 89 , 91 , 92 , 93 , 94 , 95 , 97 , 98 , 99 , 101 , 104 , 105 , 106 , 109 , 110 , 111 , 115 , 116 , 117 , 120 , 121 , 122 , 124 , 125 , 129 , 130 , 132 , 133 , 135 , 138 , 139 , 145 , 146 , 150 , 151 , 152 , 153 , 155 , 156 , 157 , 160 , 166 , 167 , 168 , 170 , 172 , 173 , 176 , 177 , 178 , 181 , 182 , 188 , 192 , 193 , 197 , 198 , 199 , 202 , 203 , 204 , 205 , 206 , 207 , 208 , 213 , 214 , 215 , 217 , 218 , 219 , 220 , 222 , 225 , 228 , 229 , 230 , 231 , 233 , 234 , 235 , 237 , 239 , 241 , 245 , 247 , 248 , 249 , 251 , 252 , 254 , 261 , 262 , 263 , 265 , 271 , 272 , 273 , 274 , 275 , 276 , 277 , 278 , 279 , 281 , 282 , 285 , 286 , 289 , 290 , 293 , 296 , 298 , 299 , 302 , 305 , 306 , 307 , 308 , 309 , 310 , 311 , 312 , 313 , 314 , 315 , 316 , 317 , 318 , 319 , 320 , 321 , 322 , 323 , 324 , 325 , 326 , 327 , 328 , 329 , 330 , 331 , 332 , 333 , 334 , 335 , 336 , 337 , 338 , 339 , 340 , 341 , 342 , 343 , 344 , 345 , 346 , 347 , 348 , 349 , 350 , 351 , 352 , 353 , 354 , 355 , 356 , 357 , 358 , 359 , 360 , 361 , 362 , 363 , 364 , 365 , 366 , 367 , 368 , 369 , 370 , 371 , 372 , 373 , 374 , 375 , 376 , 377 , 378 , 379 , 380 , 381 , 382 , 383 , 384 , 385 , 386 , 387 , 388 , 389 , 390 , 391 , 392 , 393 , 394 , 395 , 396 , 397 , 398 , 399 , 400 , 401 , 402 , 403 , 404 , 405 , 406 , 407 , 408 , 409 , 410 , 411 , 412 , 413 , 414 , 415 ,

416 , 417 , 418 , 419 , 420 , 421 , 422 , 423 , 424 , 425 , 426 , 427 , 428 , 429 , 430 , 431 , 432 , 433 , 434 , 435 , 436 , 437 , 438 , 439 , 440 , 441 , 442 , 443 , 444 , 445 , 446 , 447 , 448 , 449 , 450 , 451 , 452 , 453 , 454 , 455 , 456 , 457 , 458 , 459 , 460 , 461 , 462 , 463 , 464 , 465 , 466 , 467 , 468 , 469 , 470 , 471 , 472 , 473 , 474 , 475 , 476 , 477 , 478 , 479 , 480 , 481 , 482 , 483 , 484 , 485 , 486 , 487 , 488 , 489 , 490 , 491 , 492 , 493 , 494 , 495 , 496 , 497 , 498 , 499 , 500 , 501 , 502 , 503 , 504 , 505 , 506 , 507 , 508 , 509 , 510 , 511 , 512 , 513 , 514 , 515 , 516 , 517 , 518 , 519 , 520 , 521 , 522 , 523 , 524 , 525 , 526 , 527 , 528 , 529 , 530 , 531 , 532 , 533 , 534 , 535 , 536 , 537 , 538 , 539 , 540 , 541 , 542 , 543 , 544 , 545 , 546 , 547 , 548 , 549 , 550 , 551 , 552 , 553 , 554 , 555 , 556 , 557 , 558 , 559 , 560 , 561 , 562 , 563 , 564 , 565 , 566 , 567 , 568 , 569 , 570 , 571 , 572 , 573 , 574 , 575 , 576 , 577 , 578 , 579 , 580 , 581 , 582 , 583 , 584 , 585 , 586 , 587 , 588 , 589 , 590 , 591 , 592 , 593 , 594 , 595 , 596 , 597 , 598 , 599 , 600 , 601 , 602 , 603 , 604 , 605 , 606 , 607 , 608 , 609 , 610 , 611 , 612 , 613 , 614 , 615 , 616 , 617 , 618 , 619 , 620 , 621 , 622 , 623 , 624 , 625 , 626 , 627 , 628 , 629 , 630 , 631 , 632 , 633 , 634 , 635 , 636 , 637 , 638 , 639 , 640 , 641 , 642 , 643 , 644 , 645 , 646 , 647 , 648 , 649 , 650 , 651 , 652 , 653 , 654 , 655 , 656 , 657 , 658 , 659 , 660 , 661 , 662 , 663 , 664 , 665 , 666 , 667 , 668 , 669 , 670 , 671 , 672 , 673 , 674 , 675 , 676 , 677 , 678 , 679 , 680 , 681 , 682 , 683 , 684 , 685 , 686 , 687 , 688 , 689 , 690 , 691 , 692 , 693 , 694 , 695 , 696 , 697 , 698 , 699 , 700 , 701 , 702 , 703 , 704 , 705 , 706 , 707 , 708 , 709 , 710 , 711 , 712 , 713 , 714 , 715 , 716 , 717 , 718 , 719 , 720 , 721 , 722 , 723 , 724 , 725 , 726 , 727 , 728 , 729 , 730 , 731 , 732 , 733 , 734 , 735 , 736 , 737 , 738 , 739 , 740 , 741 , 742 , 743 , 744 , 745 , 746 , 747 , 748 , 749 , 750 , 751 , 752 , 753 , 754 , 755 , 756 , 757 , 758 , 759 , 760 , 761 , 762 , 763 , 764 , 765 , 766 , 767 , 768 , 769 , 770 , 771 , 772 , 773 , 774 , 775 , 776 , 777 , 778 , 779 , 780 , 781 , 782 , 783 , 784 , 785 , 786 , 787 , 788 , 789 , 790 , 791 , 792 , 793 , 794 , 795 , 796 , 797 , 798 , 799 , 800 , 801 , 802 , 803 , 804 , 805 , 806 , 807 , 808 , 809 , 810 , 811 , 812 , 813 , 814 , 815 , 816 , 817 , 818 , 819 , 820 , 821 , 822 , 823 , 824 , 825 , 826 , 827 , 828 , 829 , 830 , 831 , 832 , 833 , 834 , 835 , 836 , 837 , 838 , 839 , 840 , 841 , 842 , 843 , 844 , 845 , 846 , 847 , 848 , 849 , 850 , 851 , 852 , 853 , 854 , 855 , 856 , 857 , 858 , 859 , 860 , 861 , 862 , 863 , 864 , 865 , 866 , 867 , 868 , 869 , 870 , 871 , 872 , 873 , 874 , 875 , 876 , 877 , 878 , 879 , 880 , 881 , 882 , 883 , 884 , 885 , 886 , 887 , 888 , 889 , 890 , 891 , 892 , 893 , 894 , 895 , 896 , 897 , 898 , 899 , 900 , 901 , 902 , 903 , 904 , 905 , 906 , 907 , 908 , 909 , 910 , 911 , 912 , 913 , 914 , 915 , 916 , 917 , 918 , 919 , 920 , 921 , 922 , 923 , 924 , 925 , 926 , 927 , 928 , 929 , 930 , 931 , 932 , 933 , 934 , 935 , 936 , 937 , 938 , 939 , 940 , 941 , 942 , 943 , 944 , 945 , 946 , 947 , 948 , 949 , 950 , 951 , 952 , 953 , 954 , 955 , 956 , 957 , 958 , 959 , 960 , 961 , 962 , 963 , 964 , 965 , 966 , 967 , 968 , 969 , 970 , 971 , 972 , 973 , 974 , 975 , 976 , 977 , 978 , 979 , 980 , 981 , 982 , 983 , 984 , 985 , 986 , 987 , 988 , 989 , 990 , 991 , 992 , 993 , 994 , 995 , 996 , 997 , 998 , 999 , 1000 , 1001 , 1002 , 1003 , 1004 , 1005 , 1006 , 1007 , 1008 , 1009 , 1010 , 1011 , 1012 , 1013 , 1014 , 1015 , 1016 , 1017 , 1018 , 1019 , 1020 , 1021 , 1022 , 1023 , 1024 , 1025 , 1026 , 1027 , 1028 , 1029 , 1030 , 1031 , 1032 , 1033 , 1034 , 1035 , 1036 , 1037 , 1038 , 1039 ,



SET:
(-5.7378081,-6.7261898)(-5.5040044,-5.9804901)(-5.1954559,-7.2683139)(-4.6239007,-6.7752677)(-5.1039726,-7.7783371)(-5.9094735,-7.4383142)(-5.2077795,-7.2755335)(-5.5668423,-5.8461305)(-4.9876562,-6.7351985)(-5.4310054,-6.5466759)(-5.0596488,-7.412555)(-4.7627258,-6.6288991)(-4.7157247,-6.8345469)(-5.3557532,-6.1833345)(-5.4345237,-7.0232062)(-4.9337278,-7.1572124)(-4.9749228,-6.9761827)(-5.0747144,-6.7200857)(-5.2358732,-7.1027816)(-5.1278103,-7.2858295)(-4.776674,-7.0858312)(-6.2121825,-8.0213802)(-4.9413387,-6.4210905)(-5.0499548,-6.7810673)(-5.7996125,-7.8827268)(-4.7922369,-6.6833058)(-4.9171777,-6.6284922)(-4.6352562,-7.3386924)(-4.5831507,-6.9765101)(-4.9967003,-6.6172267)(-5.0764888,-7.2145599)(-4.7291797,-6.7660214)(-5.4779079,-6.3070256)(-4.8786631,-6.769412)(-4.9775755,-6.8605382)(-5.7689619,-7.4247345)(-5.0155156,-6.9578993)(-4.9454794,-7.07323)(-5.4114304,-5.9597791)(-4.7392768,-8.0932025)(-4.9144064,-7.3142177)(-5.354781,-6.056786)(-5.31057,-6.8396423)

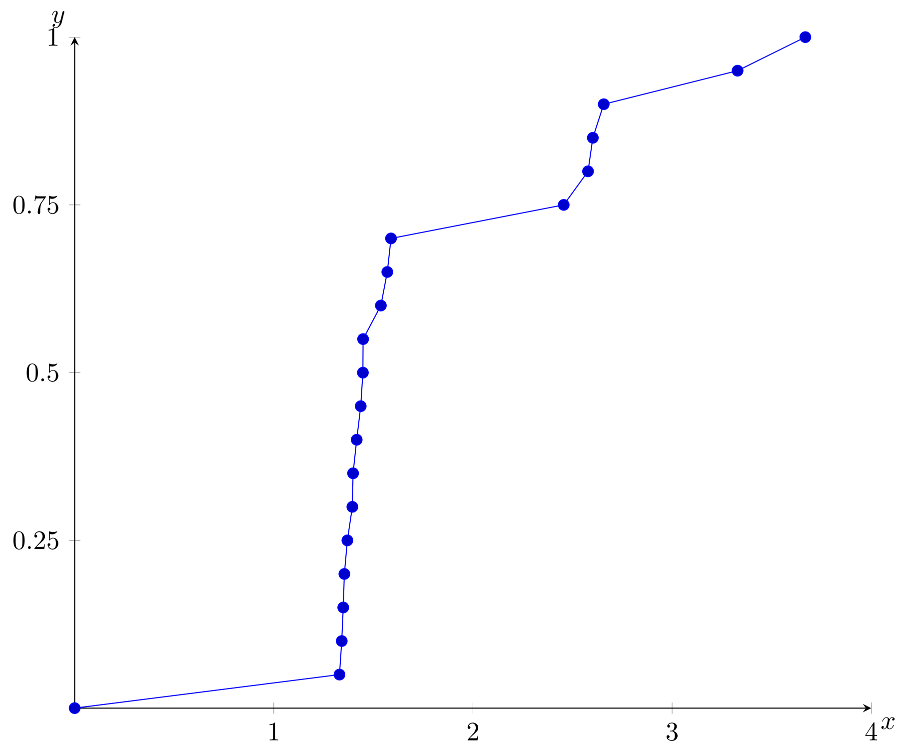


SET:
(-3.3755073,-6.8078161)(-3.7241204,-6.2411353)(-3.1538396,-6.1455886)(-4.0004525,-6.384472)(-2.8994207,-5.6713194)(-3.7348412,-6.3800545)(-3.4456156,-5.8066268)(-3.7434374,-6.0869206)(-3.8148002,-6.2660256)(-1.756621,-4.6388883)(-3.8688091,-6.302476)(-3.8893691,-6.4123887)(-3.692487,-6.4337346)(-3.8689858,-6.1779057)(-2.3395741,-5.8363895)(-4.2909707,-7.3065739)(-2.7551086,-5.3823124)(-3.5262238,-6.8995217)(-3.6305771,-6.1808633)(-3.7300596,-6.5847434)(-4.3687402,-7.5738585)(-3.7780393,-6.1308289)(-3.1543975,-5.8393938)(-2.3377491,-5.544014)(-3.7429193,-6.5328615)(-3.304779,-6.0367647)(-3.4924452,-6.8689173)(-3.4552715,-5.9550063)(-2.7275471,-5.2233346)(-3.3371386,-6.3843846)(-3.7011366,-6.0857448)(-3.7132858,-6.5209782)(-2.5486799,-5.9433799)(-2.8893272,-5.3335974)(-4.1458744,-6.421741)(-1.7405971,-5.6548925)(-2.8357246,-5.3285584)(-3.3802834,-5.7456183)(-3.3330688,-5.7062118)(-4.0444743,-6.6299313)(-3.8401947,-6.5835544)(-4.5470907,-7.1003811)(-4.1338825,-6.7728377)(-4.0755467,-6.4335519)(-2.4020712,-4.8858217)(-3.4324807,-7.2337209)(-3.7219151,-6.2338533)(-2.7460234,-5.3564866)(-4.3445077,-6.8621011)(-3.7991001,-6.5338046)(-4.4200645,-6.9371697)(-3.79842,-6.1468452)(-2.3894367,-4.9073916)(-3.6921877,-6.0451664)(-2.7935863,-

5.8288476)(-3.3581338,-6.1354949)(-3.9110232,-6.2223852)(-4.0272515,-6.3906384)(-2.9663929,-6.1638234)(-4.4348705,-6.7084391)(-3.9058857,-6.8635963)(-3.5487112,-6.1490269)(-2.9921091,-6.4791852)(-3.9797656,-6.7580299)(-1.673313,-5.6139879)(-2.892201,-5.6500006)(-4.1132262,-7.2274724)(-3.8392174,-6.9819397)(-3.1160021,-6.5080383)(-4.0744583,-6.8638336)(-3.4668627,-7.3761174)(-3.1967797,-5.8481289)(-3.5613812,-6.0371348)(-4.5631165,-7.3528581)(-4.1728637,-6.5724724)(-2.3471119,-5.6609861)(-3.3025399,-6.2729677)(-3.8137563,-6.194642)(-2.4105049,-4.9584442)(-3.672274,-6.3804859)(-3.4923589,-5.9735848)(-3.4555256,-6.3792907)(-3.2994779,-6.5800543)(-3.6813639,-6.4080051)(-3.4518845,-6.3591688)(-4.0731858,-6.4402769)(-3.3864362,-6.5561329)(-3.8204138,-6.2000687)(-3.5425328,-6.2983828)(-3.2450639,-6.9007445)(-3.2386227,-5.798026)(-3.5322929,-5.9035543)(-3.2169952,-6.0402687)(-4.3450402,-6.9170209)(-2.0494089,-5.5240076)(-3.6643534,-7.3565747)(-2.8966174,-5.757131)(-3.6443604,-6.9431494)(-3.919459,-6.3924113)(-3.3634724,-5.7461759)(-3.1061304,-5.9175077)(-3.7665645,-6.7051905)(-3.9792969,-6.8686698)(-3.876685,-6.7145541)(-3.4601547,-6.7655543)(-3.0564455,-5.7342068)(-3.4123226,-5.88231)(-3.959547,-6.419806)(0.0,-40.0)

3-means cost = 1.99
centers: (-4.01936004 -6.02337917), (0 -40), (2.49628862 4.47085919)

Run Lloyds Algorithm with C initially as the output of each run of k-Means++ above. Plot a cumu- lative density function of the 3-means cost. Also report the fraction of the trials that the subsets are the same as the input (where the input is the result of k-Means++)



14/20 times are the set the same, some times they were close, but never the same.