



< Previous	✓	✓	✓	✓	✓	✓	✓	✓	✓	Next >
------------	---	---	---	---	---	---	---	---	---	--------

Exercise 5

🔖 Bookmark this page

Exercise 5

3/3 points (graded)

As Professor Gutttag said, there are two types of people in this world: those who know programming and those who don't. To prove this once and for all, you take a random sampling of edX students and put them through a programming test. Assume that the test is entirely fair and that it reflects the exact level of skill each student has. You also ask them to fill out a small questionnaire about their experience with 6.00.2x.

You receive the results for each student as [Exam grade, Hour spent on 6.00.2x]. That is, if Alice has spent 90 hours on 6.00.2x and received a score of 74 on the exam, you will have [74, 90] as a data point.

1. Based on your initial purposes, what should you choose as k?

☒ 2

☐ 3

☐ 4



2. Should you apply scaling to this data?

☒ Yes

☐ No



You run your clustering algorithm and get two clusters:

CLUSTER 1	Exam grade	Hours spent on 6.00.2x
Mean	46.2	4.3
Variance	15.0	0.91

CLUSTER 2	Exam grade	Hours spent on 6.00.2x
Mean	84.5	60.2
Variance	5.1	6.04

Results look good – there are indeed two kinds of people, and curiously one kind seems to absolutely love 6.00.2x. However, when you rerun the code, you get the following clusters:

CLUSTER 1	Exam grade	Hours spent on 6.00.2x
Mean	12.5	2.34
Variance	0.29	0.36

CLUSTER 2	Exam grade	Hours spent on 6.00.2x
Mean	70.23	35.4
Variance	8.65	10.84

3. You don't know what to believe (and, indeed, there's no reason for you to choose one over another). What can you do to fix this and get a stable result?

☒ Let k = 3

Calculator



Add more

< Previous

Next >



Add "student name length" as a feature



© All Rights Reserved



edX

- [About](#)
- [Affiliates](#)
- [edX for Business](#)
- [Open edX](#)
- [Careers](#)
- [News](#)

Legal

- [Terms of Service & Honor Code](#)
- [Privacy Policy](#)
- [Accessibility Policy](#)
- [Trademark Policy](#)
- [Sitemap](#)

Connect

- [Blog](#)
- [Contact Us](#)
- [Help Center](#)
- [Security](#)
- [Media Kit](#)



© 2022 edX LLC. All rights reserved.
深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)