



Microsoft: DAT210x Programming with Python for Data Science



Bookmarks

- ▶ Start Here
- ▶ 1. The Big Picture
- ▶ 2. Data And Features
- ▶ 3. Exploring Data
- ▼ 4. Transforming Data

Lecture: Transformations

Lecture: PCA

Quiz



Lab: PCA

Lab



Lecture: Isomap

Quiz



Lab: Isomap

Lab



Lecture: Data Cleansing

Quiz



4. Transforming Data > Lecture: Isomap > Knowledge Checks



Bookmark

Review Question 1

(1/1 point)

Which of the following explanations of isomap is true?

- ☐ Isomap can be used as a powerful noise removal tool, since a smooth manifold is created by "short circuiting" the nearest neighbor map when calculating distances
- ☐ Isomap is usually faster than PCA because it's quicker to compute a nearest neighbor map than to do matrix decomposition
- ☒ A one sentence summary of isomap's implementation is that at its core, it is essentially a node distance map that has been fed into a special type of PCA ✓
- ☐ Isomap will not function without a completely accurate distance metric, since it needs to know the precise distance to every single sample, including distant ones

EXPLANATION

Dive Deeper

► 5. Data Modeling

One of isomap's greatest weaknesses is that noisy data might short circuiting the actual geodesic path. In such cases, isomap will prefer the noisy path to the actual path and produce an incorrectly warped mapping.

Isomap is slower than PCA because isomap essentially implements a multi-dimensional scaling (similar to PCA) through projection; but in addition to that, it also has to calculate the nearest neighbor map.

As mentioned in the lecture, even if the distance metric isn't 100% accurate, isomap can still function reasonably. Particularly for distant nodes, they won't be included in the *nearest* neighbors list to start with.

You have used 2 of 2 submissions

Review Question 2

(1/1 point)

Isomap is most beneficial...

- ☐ When your data lacks an inherent manifold
- ☒ When a non-linear, geometric structure is expressed in your data ✓
- ☐ When you are uncertain how many samples are needed to capture the underlying nature of your data

- ☐ When your high dimensionality data has a hidden, linear relationship expressed within it

You have used 1 of 2 submissions

© All Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

