



## 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 > Dive Deeper > Further Reading



Bookmark

## Dive Deeper

Congratulations on making it this far! In the previous module, you directly explored your data using many visualizations. This time you learned how to take complex datasets and simplify them using two popular methods: keeping the most variant set of orthogonal components, and manifold multi-dimensional scaling of your sample's distance map.

After that you learned about a way errors can creep into your dataset and potential methods of handling that. Keep all these techniques fresh in your toolbox by being sure you record some notes about them in the transforming section of your course map. Remember, there is no hard order you must stick to while applying these methods. Try something out, visualize your data, then continue experimenting until you get your desired results.

Below, as usually, we've included some added details about the techniques you just studied in case you're interested in further broadening your knowledge, and taking it to the next level! For instance, regarding the isomap lab, you learned what each parameter does and how isomap works; but one thing you might be wondering from our examples is how *interpolation* between video frames is performed using isomap. You know it's impossible to `.inverse_transform()` an isomap manifold, so how in the world might you go about interpolating data in your original feature space / dataset? Check below!

## PCA

- Interpreting PCA

**Dive Deeper**

## ► 5. Data Modeling

- Another Method for Interpreting PCA
- Interactive PCA Demo
- PCA on Binary Data
- The Best Explanation of the Math Behind PCA You'll Ever Read
- RandomizedPCA
- Correlation or Covariance?

**Isomap**

- The Historic Stanford Isomap Paper, By Josh Tenenbaum
- Lower Dimensional Embedding
- Manifold Learning
- Interpolating Images Between Video Frames Using Non-Linear Dimensionality Reduction
- Cambridge Hand Gesture Data set

**Climate Change within the Continental United States**

- NOAA U.S. Climate Divisional Dataset
- Interactive Discovery Tool For Comparing Raw Historical Data With the 'Enhanced' Data (Flash Based)
- Control Types

© 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.

