



KEY POINTS

The data science process involves:

1. Data selection.
2. Preprocessing.
3. Transformation.
4. Data Mining.
5. Interpretation and evaluation.

It is an iterative process in which some, or all, steps may be repeated.

FURTHER READING

Note: These links take you to external sites outside of the edX course.

- Computing Community Consortium Big Data Whitepaper: <http://cra.org/ccc/wp-content/uploads/sites/2/2015/05/bigdatawhitepaper.pdf>
- From Data Mining to Knowledge Discovery: <http://www.aaai.org/ojs/index.php/aimagazine/article/view/1230>
- CRISP-DM 1.0: http://spss.ch/upload/1107356429_CrispDM1.0.pdf

© 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.

POWERED BY
OPENedX

