

**Microsoft: DAT203x Data Science and Machine Learning Essentials**

- ▶ Before You Start
- ▶ Module 1: Introduction and Data Science Theory
- ▶ Module 2: Working with Data
- ▶ Module 3: Visualization, and Building and Evaluating Models
- ▶ Module 4: Regression, Classification, and Unsupervised Learning
- ▼ **Module 5: Recommenders and Publishing Your Work**

Chapter 19:
Recommendation
Models

Lab 5A: Working
with
Recommendation
Models

Chapter 20:
Introduction to
Jupyter Notebooks
in Azure ML

QUESTION 1 (1/1 point)

You create an experiment and implement a data flow to cleanse the data. You need to split the data into two sets for training and testing a Matchbox Recommender module in such a way that the user-item pairs are evenly divided into training and test sets.

What should you do to split the data?

☐ Add a Split module and configure it to use the Split Rows splitting mode with the fraction of rows in the first output set to 0.5.

☐ Connect the source dataset directly to the Train Matchbox Recommender module without splitting the data.

☐ Add a Project Columns module and include only the UserID and ItemID columns.

☒ Add a Split module and configure it to use the Recommender Split splitting mode. ✓


EXPLANATION

To split data for recommendation, use the Recommender Split splitting mode of the Split module. This evenly distributes user-item pairs into the training and test sets.

You have used 2 of 2 submissions

Chapter 21:
Publishing Azure
ML Models

Lab 5B: Publishing
Models in Azure
ML

Module 5 Review
Homework due Oct 30,
2015 at 00:00 UTC 

► Final Exam

© 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

