A-Z Machine Learning using Azure Machine Learning (AzureML)

Hands on AzureML: From Azure Machine Learning Introduction to Advance Machine Learning Algorithms. No Coding Required.

BEST SELLER ★★★★★ 4.3 (215 ratings)   1,597 students enrolled

Created by Jitesh Khurkhuriya   Last updated 3/2018   💬 English   🄲🄲 English

Gift This Course

Preview This Course

Section 03
Data Processing with Azure ML

Data Input Output To Azure Workspace

Enter Data Manually

Upload a Dataset

Data Format Conversion

Import Data

Unpacked Zipped Dataset

Data Manipulation

Add columns

Add Rows

Apply SQL Transformation

Remove Duplicate row

Select Columns in a Dataset
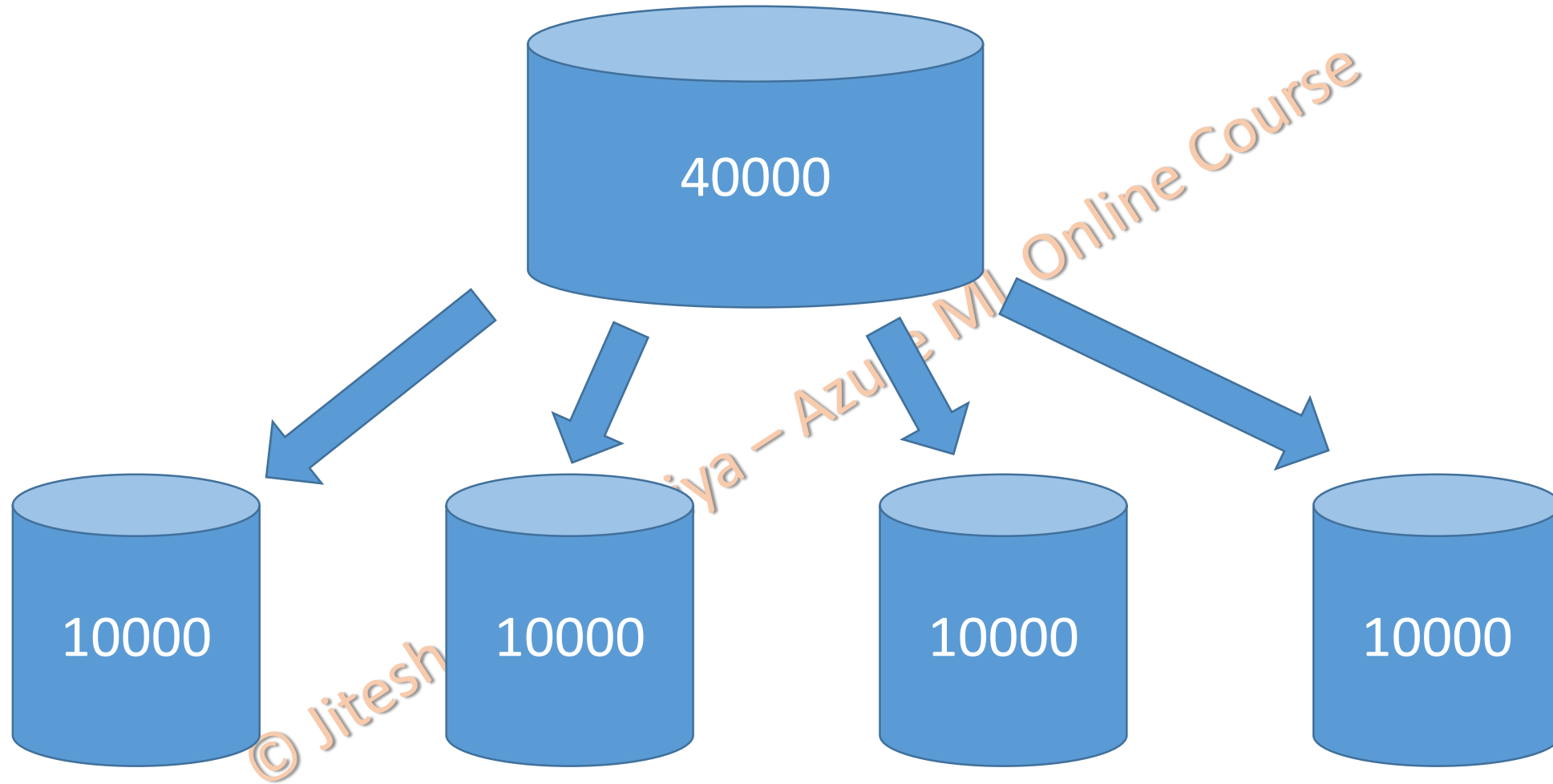
Edit Metadata

Clean Missing Data

Sample and Split
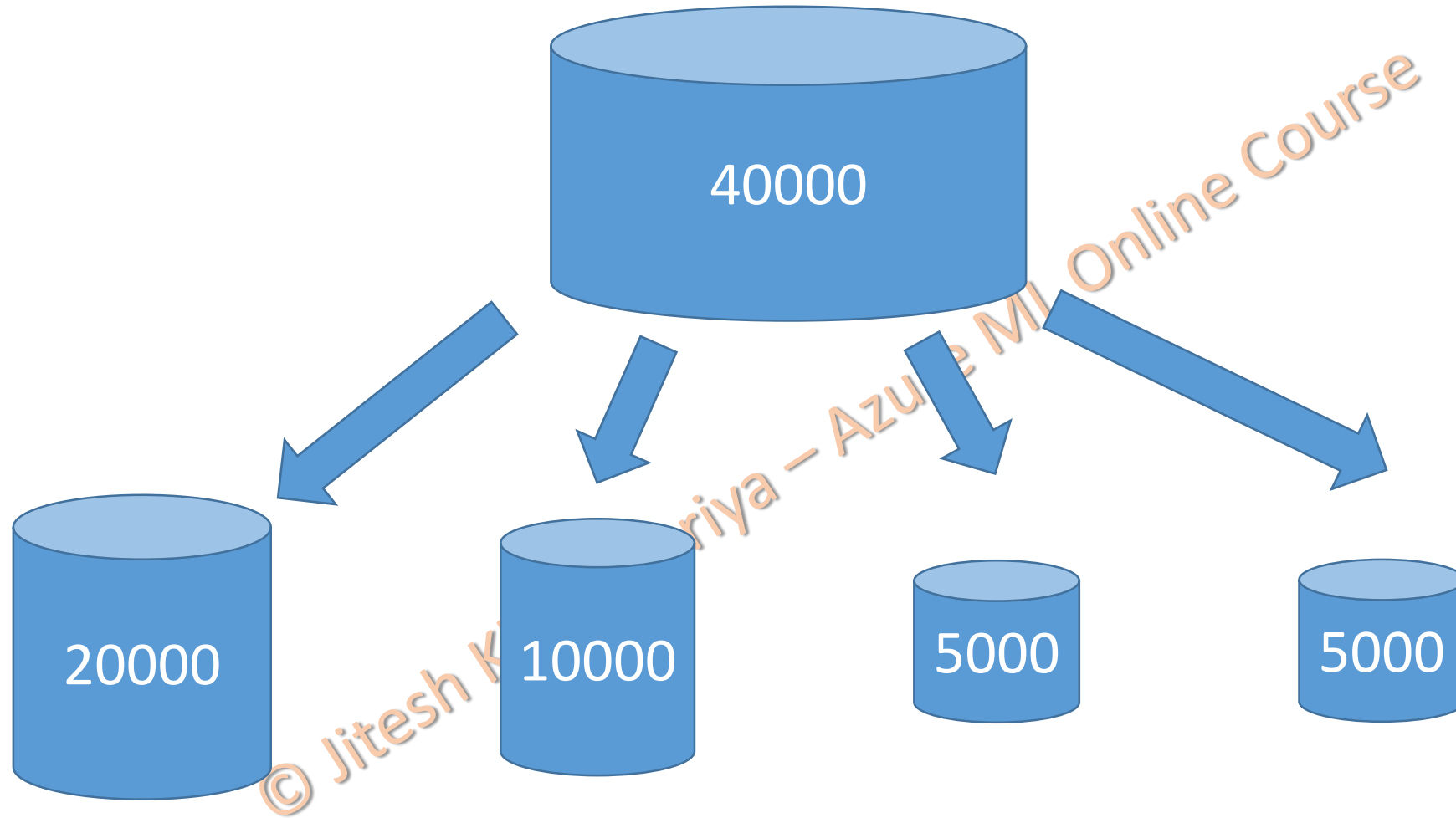
Partition and Sample

Split Data

Partition and Sample

Dividing your data into multiple subsections of the same size.

Partition and Sample

Dividing your data into multiple subsections of different sizes.

Why to partition?

- The input dataset is huge and have too many observations

- Partition the data to create train and test datasets

- Create different bins of data for cross validation of your results

- Create a random sample of observations or a more balanced dataset for a particular column

Stratification

Rating	Number of Observations
Rating 3	20
Rating 5	5
Total	25

Sampling rate 0.4 or 40%   No Stratification

Rating	Number of Observations
Rating 3	10
Rating 5	0
Total	10

Rating	Number of Observations
Rating 3	5
Rating 5	5
Total	10

Rating	Number of Observations
Rating 3	7
Rating 5	3
Total	10

Stratification

Rating	Number of Observations
Rating 3	20 x 0.4
Rating 5	5 x 0.4
Total	25

Sampling rate 0.4 or 40%   With Stratification

Rating	Number of Observations
Rating 3	8
Rating 5	2
Total	10

Thank You and Have a Great Time !