OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 17 - API to Pa...
- Day 18 - Web Scra...
- Day 19 - Understa...
- Day 20 - Univariate...
- Day 21 - Bivariate...
- Day 22 - Pandas Pr...
- Day 23 - Feature E...
- Day 24 - Standardi...
- Day 25 - Normaliza...
- Day 26 - Ordinal E...
- Day 27 - One Hot E...
- Day 28 - ColumnTr...
- Day 29 - Pipelines
- Day 30 - Function T...
- Day 31 - Power Tra...
- Day 32 - Discrtizati...
- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...

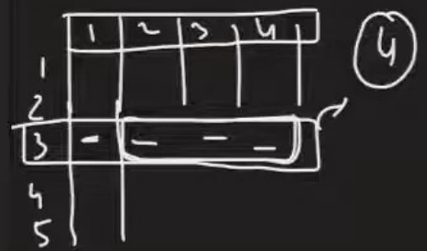
Add section

Handling Missing Data
Complete Case Analysis
Assumption For CCA
Advantage/Disadvantage
When to use CCA?
Example

Add page

Handling Missing Data

Thursday, April 22, 2021 1:03 PM



OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 17 - API to Pa...
- Day 18 - Web Scra...
- Day 19 - Understa...
- Day 20 - Univariate...
- Day 21 - Bivariate...
- Day 22 - Pandas Pr...
- Day 23 - Feature E...
- Day 24 - Standardi...
- Day 25 - Normaliza...
- Day 26 - Ordinal E...
- Day 27 - One Hot E...
- Day 28 - ColumnTr...
- Day 29 - Pipelines
- Day 30 - Function T...
- Day 31 - Power Tra...
- Day 32 - Discrtizati...
- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...

Add section Add page

Handling Missing Data

Complete Case Analysis

Assumption For CCA

Advantage/Disadvantage

When to use CCA?

Example

Complete Case Analysis

Thursday, April 22, 2021 12:55 PM

Complete-case analysis (CCA), also called "list-wise deletion" of cases, consists in **discarding** observations where values in **any** of the variables are missing.

Complete Case Analysis means literally analyzing only those observations for which there is information in **all** of the variables in the dataset.

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 17 - API to Pa...
- Day 18 - Web Scra...
- Day 19 - Understa...
- Day 20 - Univariate...
- Day 21 - Bivariate...
- Day 22 - Pandas Pr...
- Day 23 - Feature E...
- Day 24 - Standardi...
- Day 25 - Normaliza...
- Day 26 - Ordinal E...
- Day 27 - One Hot E...
- Day 28 - ColumnTr...
- Day 29 - Pipelines
- Day 30 - Function T...
- Day 31 - Power Tra...
- Day 32 - Discrtizati...
- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...

Add section

Handling Missing Data

Complete Case Analysis

Assumption For CCA

Advantage/Disadvantage

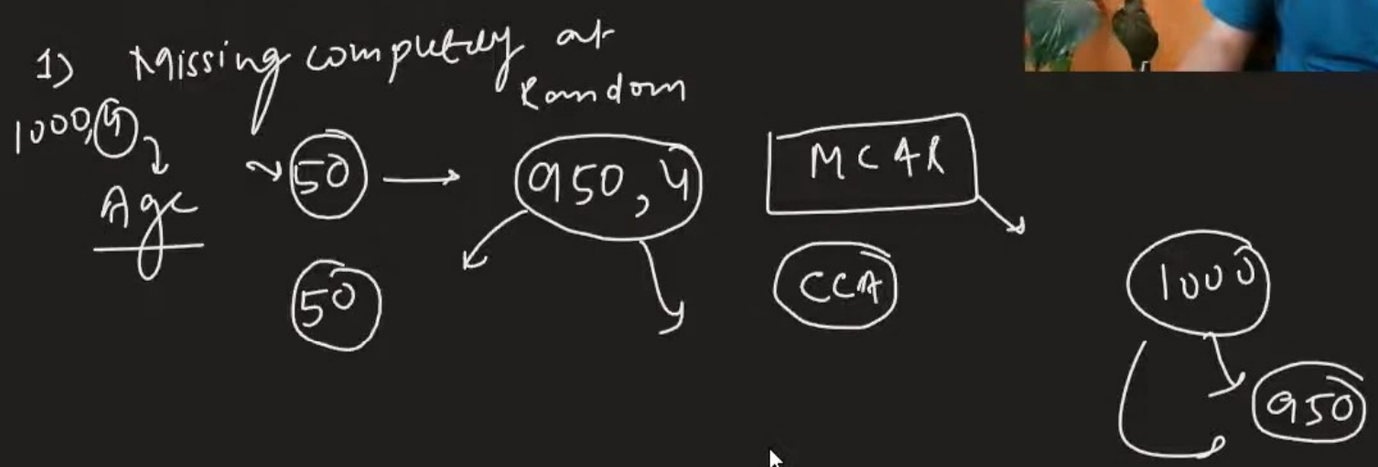
When to use CCA?

Example

Add page

Assumption For CCA

Thursday, April 22, 2021 12:58 PM



Advantage/Disadvantage

Thursday, April 22, 2021 12:59 PM

Advantage

1. Easy to implement as no data manipulation required
2. Preserves variable distribution (if data is MCAR, then the distribution of the variables of the reduced dataset should match the distribution in the original dataset

Disadvantage

1. It can exclude a large fraction of the original dataset (if missing data is abundant)
2. Excluded observations could be informative for the analysis (if data is not missing at random)
3. When using our models in production, the model will not know how to handle missing data

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

Day 17 - API to Pa...
Day 18 - Web Scra...
Day 19 - Understa...
Day 20 - Univariate...
Day 21 - Bivariate...
Day 22 - Pandas Pr...
Day 23 - Feature E...
Day 24 - Standardi...
Day 25 - Normaliza...
Day 26 - Ordinal E...
Day 27 - One Hot E...
Day 28 - ColumnTr...
Day 29 - Pipelines
Day 30 - Function T...
Day 31 - Power Tra...
Day 32 - Discrtizati...
Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...

Add section

Handling Missing Data
Complete Case Analysis
Assumption For CCA
Advantage/Disadvantage
When to use CCA?
Example

Add page

When to use CCA?

Thursday, April 22, 2021 1:02 PM