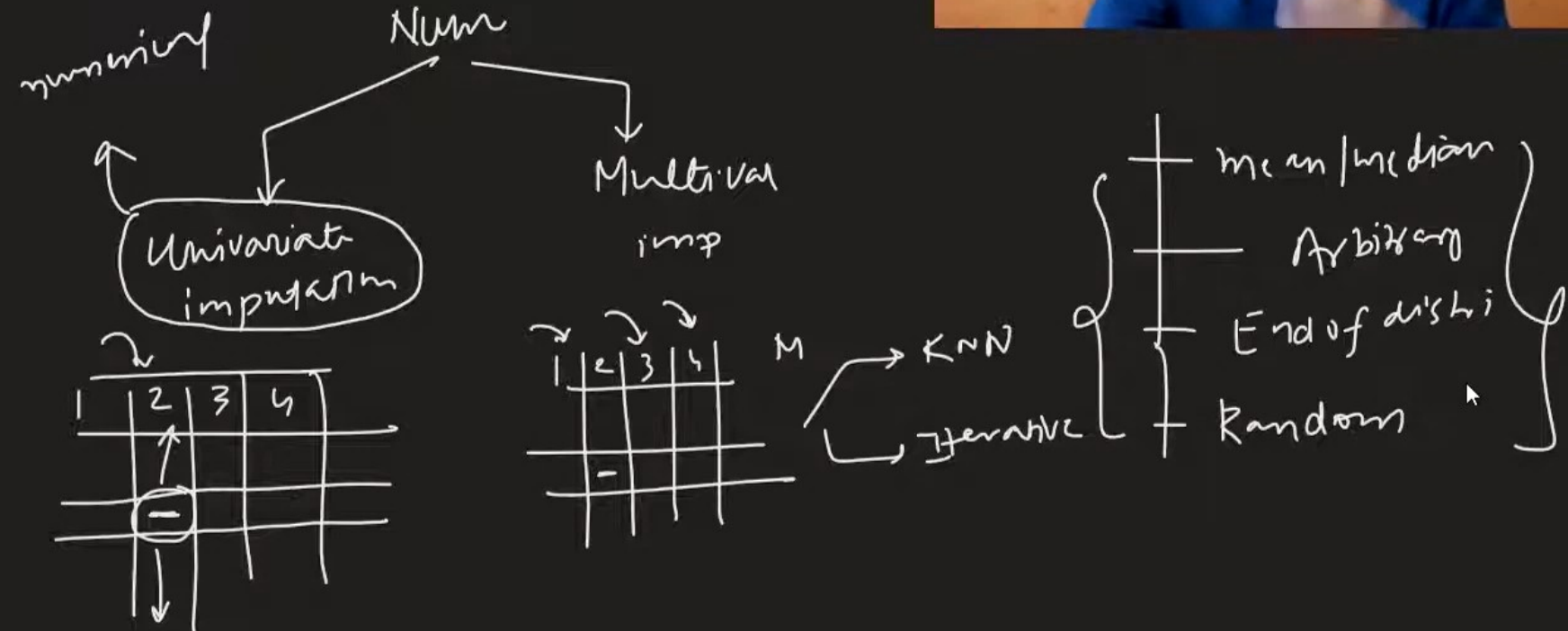




1. Handling Missing Numerical Data

Friday, April 23, 2021 11:28 AM



100 Days of ML

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...

Day 36 - Handling...

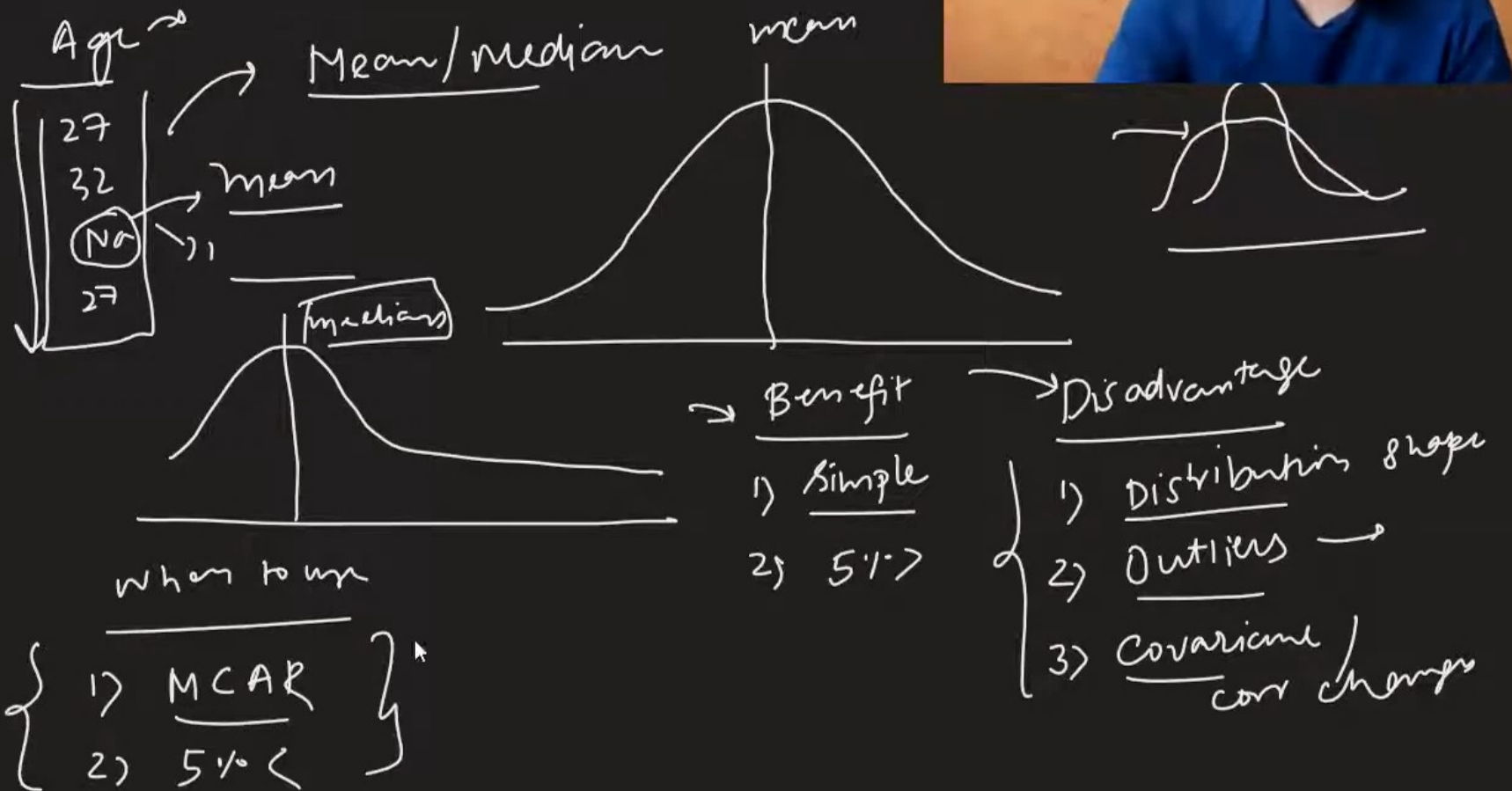
+ Add section

+ Add page



2. Mean/Median Imputation

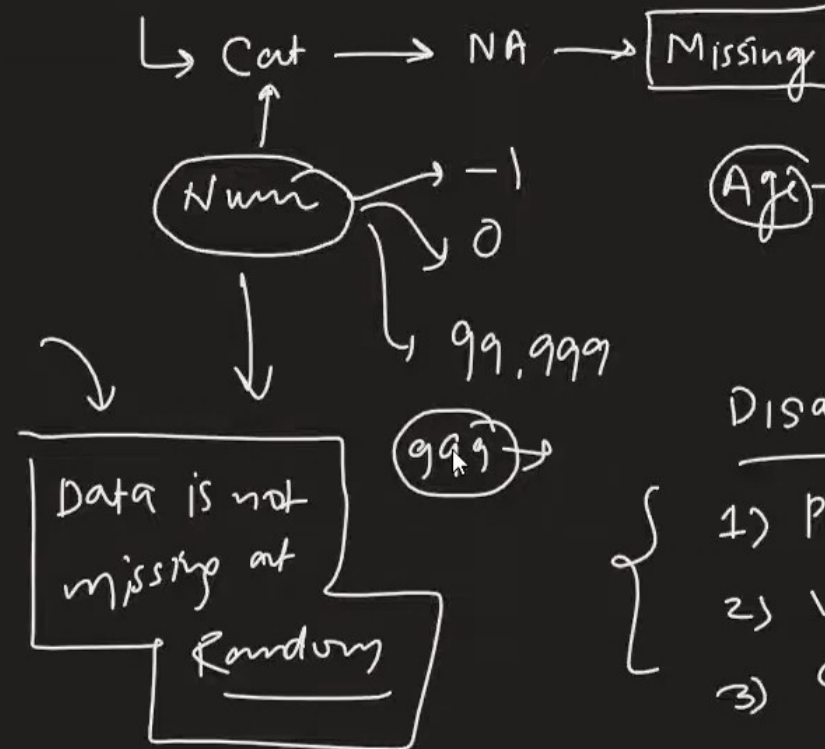
Friday, April 23, 2021 11:31 AM





3. Arbitrary Value Imputation

Friday, April 23, 2021 11:32 AM



Age - 99
-1

Benefit
↳ easy to apply

Disadvantage

- 1) PDF distort
- 2) variance
- 3) covar changes

100 Days of ML

- 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...
- Day 36 - Handling...

+ Add section

+ Add page

100 Days of ML

🔍

🕒

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...

Day 36 - Handling...

1. Handling Missing Numerical Data

2. Mean/Median Imputation

3. Arbitrary Value Imputation

4. End of Distribution Imputation

5. Random Sample Imputation

6. Automatically select best imput...

+ Add section

+ Add page

4. End of Distribution Imputation

Friday, April 23, 2021 11:32 AM



Normality
↳ (mean + 3σ)
mean - 3σ

Skewed (IQR proximity)
↳ Q1 - 1.5 IQR
Q3 + 1.5 IQR
25th Q1
75th Q3

