

← → OneNote for Windows 10

Home Insert Draw View Help

↶ ↷ AI + ⇅

📁 100 Days of ML ▾

🔍 🕒

- 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...
- Day 37 - Handling...
- Day38-Missing Indi...

+ Add section + Add page

Topics to be covered

Monday, April 26, 2021 3:38 PM

⑥

```
graph TD; Root((6)) --> CCA[CCA]; Root --> Numerical[Numerical]; Root --> Categorical[Categorical]; Root --> RandomImputation[Random Imputation]; Root --> Interactive[Interactive]; RandomImputation --> Random[Random]; RandomImputation --> MissingIndicator[Missing Indicator]; RandomImputation --> AutomaticSelect[Automatic select];
```



- 100 Days of ML
- 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...
- Day 37 - Handling...
- Day38-Missing Indi...

- Topics to be covered
- Random Imputation
- Missing Indicator
- Automatically select value for Imp...

Missing Indicator

Monday, April 26, 2021 3:41 PM

Age	Female	Age-Na
27	32	F
41	35	F
Na	41	T
62	32	F

model

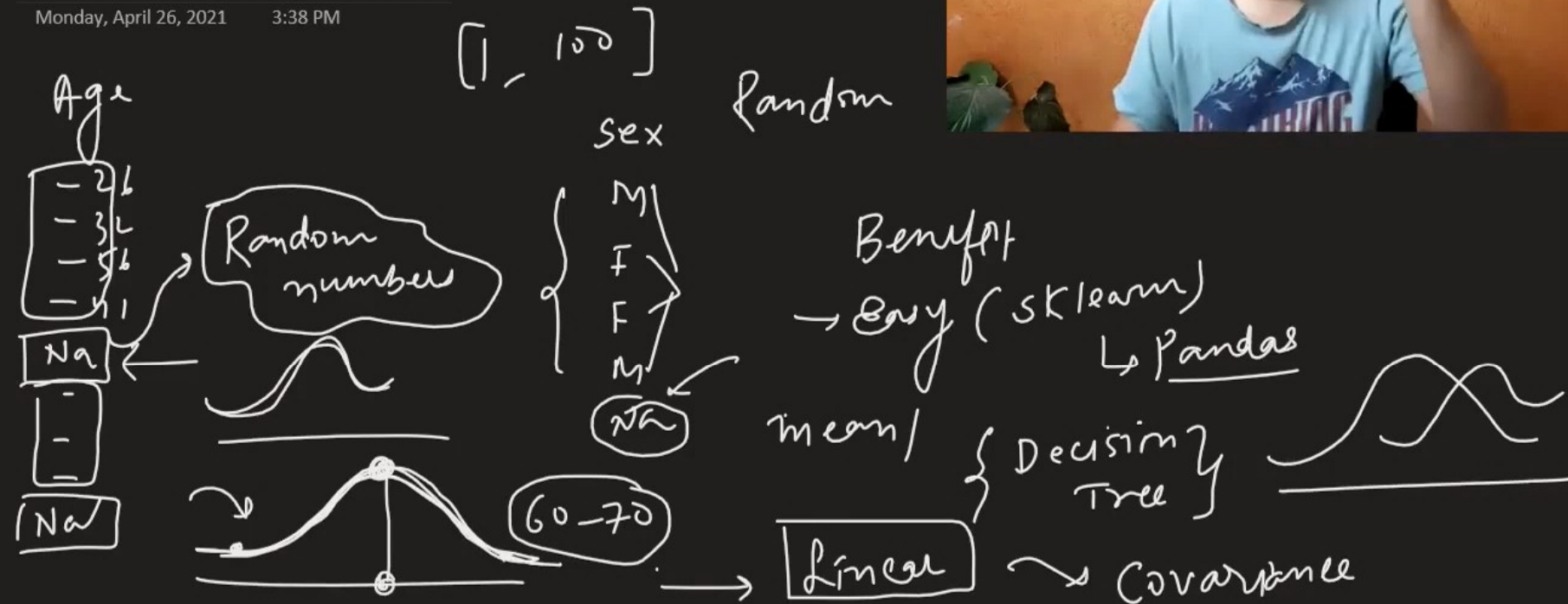
KDD



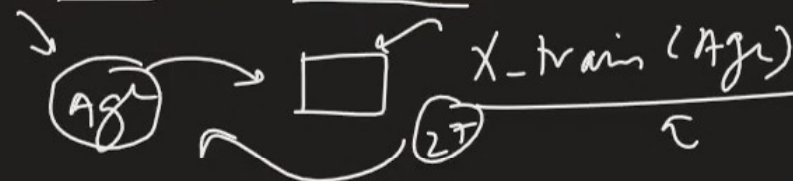


## Random Imputation

Monday, April 26, 2021 3:38 PM



1. Preserves the variance of the variable (But Why?)
2. Memory heavy for deployment, as we need to store the original training set to extract values from and replace the NA in coming observations
3. Well suited for linear models as it does not distort the distribution, regardless of the % of NA



### 100 Days of ML

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...

Day 37 - Handling...

Day38-Missing Indi...

+ Add section

+ Add page