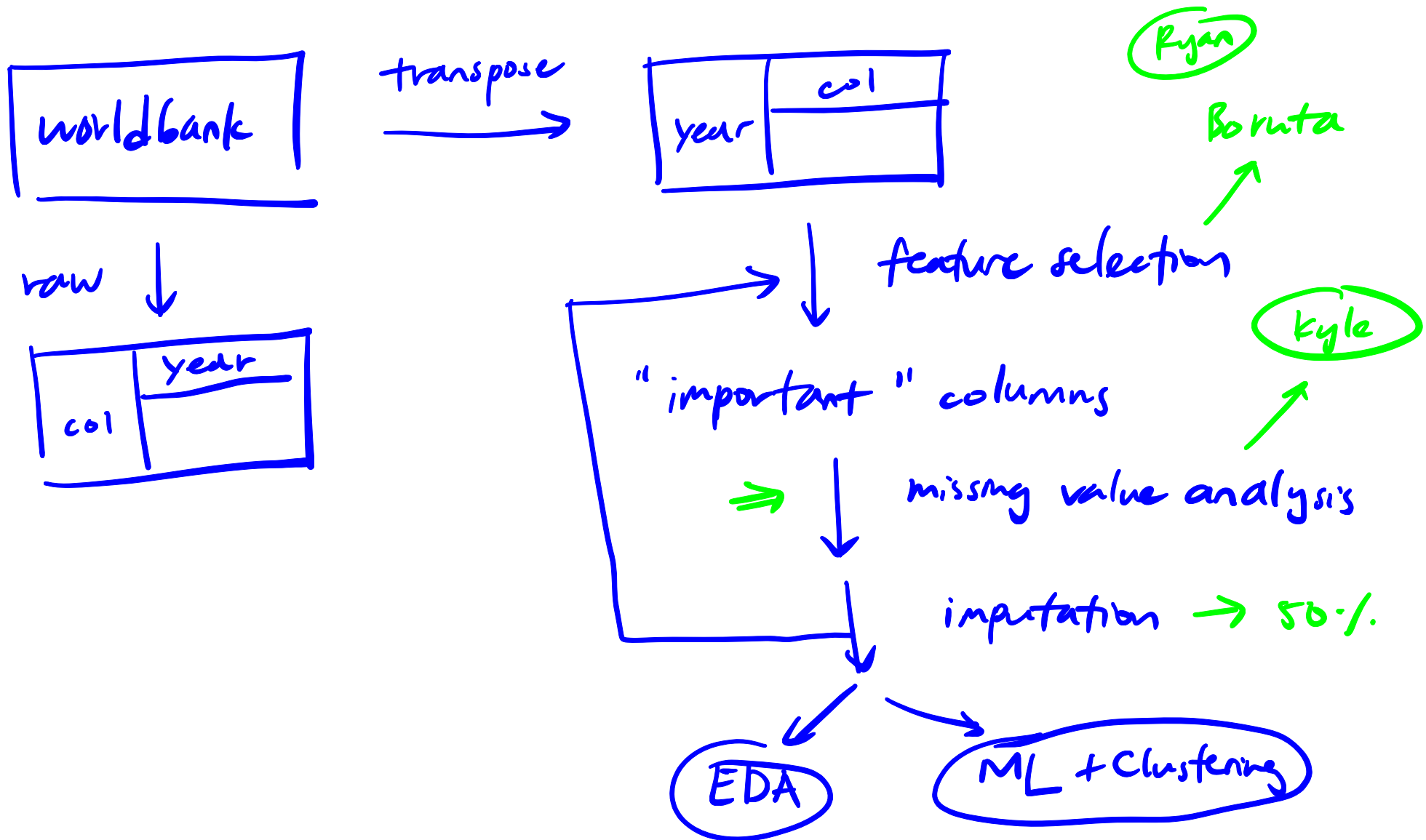


What are the different groups of CO_2 emission countries?



New Dataflow

worldbank

col	year

transpose
→

year	col



remove noisy country



drop col with >50% missing

fill NA with
median



feature selection
(BORUTA)

columns to keep



"new"



Multivariate
imputation

"clean"



EDA

ML + Clustering

