Notes and background

## might be super slow for bigger data sets