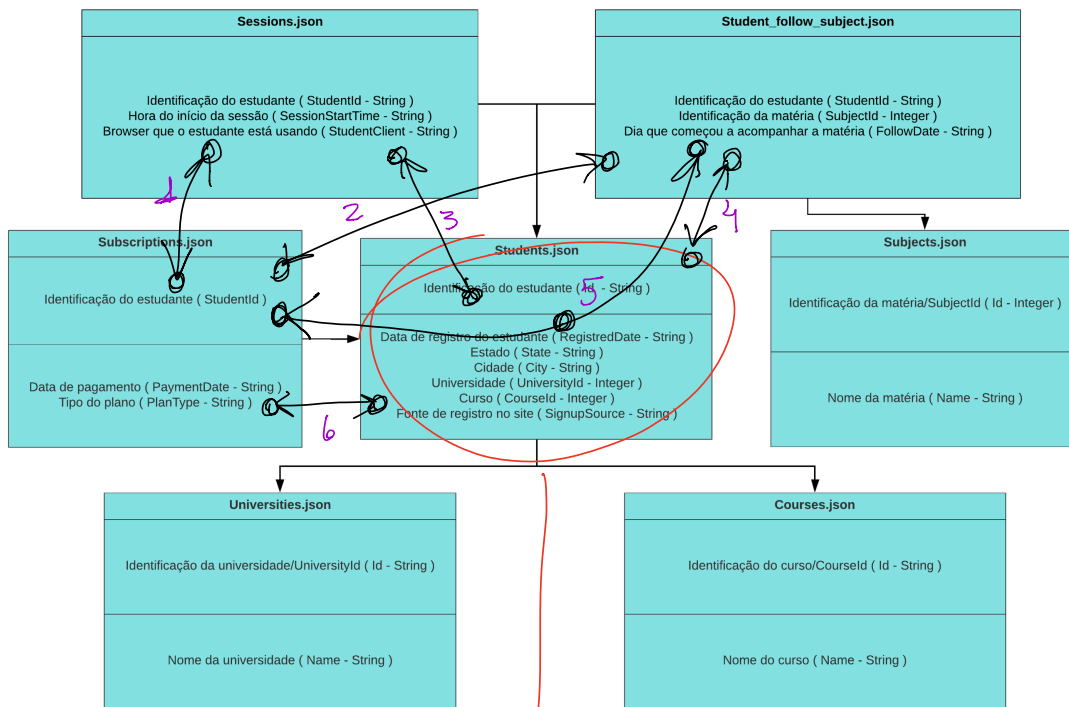


6 BD's novos para responder as perguntas



Depende do usuário:
 tem o ~~not~~ Null em
 State e City, porque?
 Olha a plataforma e
 tenta pensar.

Das variáveis cadastradas no tempo X,
 quais as ^{principais} características dos que mais entraram
 no site?

60000 cadastrados pra 283941 entradas

Banco de Dados de 2017 provavelmente
pois o último iOS registrado era iOS 11

Talvez de para tirar alguma relação dos
aparelhos mais frequentes e tipo de registro?

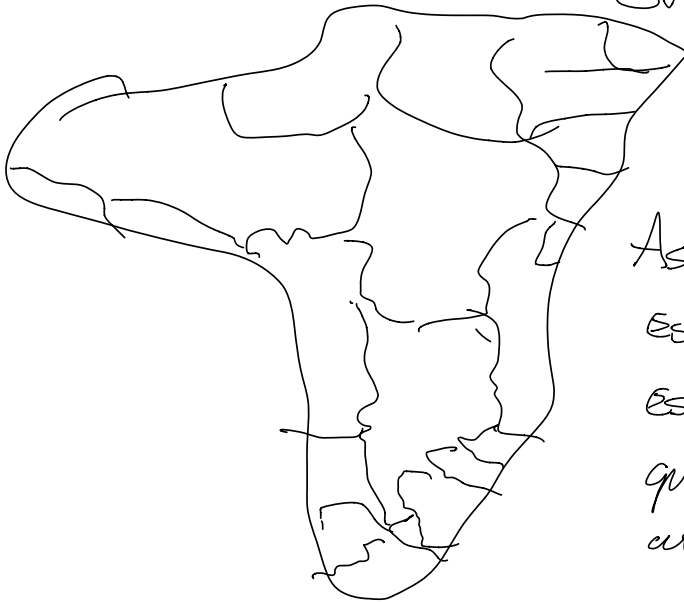
↑
Face/Google/E-mail

X stud segue Y matéria

W pessoas da uni R seguem matéria Y

Com Boleh

Brasil



As cores de cada
Estado ficam mais
escuras em rel. a
quantidade de usuários
ativos naquele MM/AAAA



talvez tenha como derivar o estado e
cidade dos usuários para imputar os dados
faltantes pela universidade, data de registro e
tipo de assinatura que entrou

no site (pg aí delimita)
a data das provas
da instituição

Tenho que ver as características dos usuários que não batam esses dados e pensar como poderia fazer pra baterem.

Se é um Data Pipeline, seria conveniente bater o wrapper de Numba para tornar o processo mais eficiente.

Cosas para prever:

- Tendência de uso de aparelhos
- Qual faculdade próximos meses vai ter mais assinantes?
- Qual faculdade próximos meses vai ter mais visualizações/acessos?

— Ver a correlação faculdade x visu, faculdade x acessos contra estado x visu e estado x acessos, avaliar as mais fortes, e validar como fazer uma

análise por ¹plataformas é mais ^{0 0}precisa.

- Pegar dados de quantos usuários no Brasil de celular usam iOS / Android, deixar ^{proporcional} para avaliar a performance ^(inter) do app de um usuário de iOS e Android deveria ser diferente em algum sentido, pq se normalizar ambas e o tráfego em uma por ~~not~~ diff do outro é possível que tenha algo aí.

- Quais assuntos mais consumidos?

Qual a tendência a aumentar?

Das 10 principais universidades, quais os assuntos mais ~~com~~ acessados, seguidos?

Os alunos de quais universidades tendem a assinar o pacote premium, e outra / quais assuntos os alunos que assinam o pacote tendem a consumir?

Conversão de dados em uma versão
que uma rede neural pode processar;
(Introduction to DL for Tabular Data, part 1)

fez um BD para isso

fez um ^{cópia} BD que faz imputação de dados e
valida a qualidade desse novo BD

Para sugestões para análises futuras, seria
tipo usar MLN para extração de conhecimento
nos BD.