**Document on merged data**

**weblog\_p1p2: combined weblog df for p1 and p2**

datetime, method, url, protocol, status, byte\_size, referrer, Session\_ID, user\_agent, category

(same as previously)

**df\_mouse\_p1p2: combined mouse df for p1 and p2**

session\_id: primary key

mousemove\_times: the timestamp of when each mouse move was performed

mousemove\_type: m for movement, c for click

mousemove\_behavior: coordinates in the form (x, y)

label: moderate\_bot, advanced\_bot, or human

mousemove\_client\_height\_width: (x, y) for those rows from phase 2, NaN for rows from phase 1

**mouse\_weblog\_merge: mouse df with some columns from weblog df**

session\_id

mousemove\_times

mousemove\_type

mousemove\_behavior

label

mousemove\_client\_height\_width

status\_mode: mode value of status for that session\_id

user\_agent

avg\_byte\_size: average byte size for that session\_id

\* this df is not saved but code is in the ipynb under the text header “mouse\_weblog\_merge: mouse df with some columns from weblog df” all the way below

\* not enough ram to run in collab or jupyter notebook