MEMORANDUM
To: BDA Students interested in the Sony Playstation Network project
From: Prof. Roger Bohn
Subject: Preliminary information about Sony's data
Date: April 19, 2018

This project is exciting, but it is going to require a high degree of adaptability and self-teaching. If you don't have the time, energy, or mental flexibility for this, *choose a different project.* You will save yourself some blood, sweat, toil, and tears. On the other hand if you are good at digging into online resources and figuring things out for yourself, it's an opportunity to work with real data that somebody cares about and has not had time to analyze thoroughly.

Our Sony alumni, Blake, gave me some partial documentation. If you are still interested in this project, **get busy immediately,** without waiting for the true data to emerge. You can find archived "clickstream" data on the web from other sites, and begin to explore what is in it. They use Adobe's system for analyzing click data, whose documentation is available at https://marketing.adobe.com/resources/help/en_US/reference/datafeeds-reference.html
There is probably generic "clickstream" data (think Apache server logs) either in well-known sample data sets, or included with some online learning tutorials, that would be closest to what we have.

Here is Blake's list of some of the most interesting issues and variables.

What do people want to do with PSN on web + mobile?
Which activities should we support?
What kinds of devices should we support?
Do people buy games?
Are people engaged with social features?

A few variables in the data dictionary
campaign
pageName
channel
mobiledevice

browser

evar49

evar40

evar16

post_custID

evar75

first_hit_time_gmt

Tip: Use "post_" versions for "downstream attribution"

Gotchas

Variables can be duplicate - no meaningful correlation

eg evar40 and post_evar40

Some IDs could look numeric but should be treated as categorical

Very large volumes of data

Do plenty of exploratory analysis first

Test your code on smaller samples ~~if necessary~~

Action: Submit a new proposal by **Sunday noon** at the latest, demonstrating that you have done some independent research to better understand issues like clickstream mining, Sony mobile games, and the Playstation Network web site. Sign up for PSN to get a sense of what they are selling.

If you decide not to pursue this opportunity, submit another proposal immediately. Follow the two Projects Assignments for information on what to provide.