+--------------------------+

| PROJECT INSTRUCTIONS |

+--------------------------+

============

Introduction

============

A company working with online data wishes to gain insight on the intentions of visitors in order to optimize the purchases these visitors make and maximize sales. To optimize purchases, the company needs to personalize the contact option for the website. Therefore, the company needs to identify who of the visitors they should offer to 'help', what to offer each, where on the site to do so, and when to show the offer. As a starting point for answering these questions, your task is to predict whether a visitor is going to accept/reject an offer for help at this time. The enclosed data file describes online behavior from a website with several different contact offers. Many features are collected in general and are described below under 'Metadata'.

===========

What to do?

===========

\* Use the data to create a predictive model for predicting whether a visitor's online behavior/profile will result in a contact or not, ie predict the last column 'contact'.

\* Evaluate the performance of the model and draw conclusions. Also, briefly state why you have chosen a particular approach, features, etc.

\* Try to gain additional insight on what makes visitors accept or reject a contact offer (eg using descriptive analytics and statistics). State any interesting insights you may have found (as comments).

\* How do you propose addressing the other questions on what to offer, where, and when to show the offer? State in comments.

\* How would you assess if your entire approach, once conceived, has an actual impact on business? State in comments.

\* Submit your python source code in 1 plain text file.

!Make sure your code runs: your model will be evaluated on a test set which has not been shared with you.

Questions? This is a project: there is no single \*correct\* solution or path; do not ask your instructors for help beyond this README file.

=============================

Metadata (ie data about data)

=============================

\* File name: data.csv

\* Description: logged features on online behavior of visitors and whether they accepted/rejected a contact offer. The offer may be in the form of a phone number, a chatbox, etc. A visitor may receive more than one contact offer during a visit. If you find the data not perfectly 'clean' or consistent, you will have to work around data quality issues which is almost always the case in real life. Values coded as NA are missing (not available).

\* Fields:

- uid: unique user id

- landing\_page: the first page viewed by the user in this session (anonymized URL)

- traffic\_source: where the visitor came from when this session started. Can be one of the following:

1. Search: this visitor used a search engine and clicked on a regular link

2. SEM: this visitor used a search engine and clicked on an advertisement

3. Affiliate: this visitor clicked an advertisement on a site which is not a search engine

- first\_session\_start\_time: the first time that this visitor was observed on the website

- total\_visits: the total number of times the visitor has visited the website

- previous\_session\_start\_time: the last time the visitor visited the website; NA if there was no previous visit

- current\_session\_start\_time: the time that this visit started

- display\_time: the time at which the contact offer was made

- clicks\_before\_this\_display: the number of clicks (pageviews) that this visitor made during this visit before receiving the contact offer

- previous\_displays\_this\_session: the number of previous contact offers made during this session

- previous\_display\_time: the timestamp of the previous contact offer; NA if there were no previous contact offers during the session

- current\_url: the (anonymized) URL of the page the contact offer was made on

- country: the country that this visitor comes from (based on IP address); the country names have been anonymized

- message\_type: an ID that identifies the type of content that was used for the offer, eg: a popup or a small phone number in the footer

- response\_promise: some offers promise a response time such as 'we will call you within 1 minute' or 'we will call you later today'.

- display\_reason: an ID corresponding to the reason that the offer was made: this may be because of user behavior during the session or it may just be that certain pages always have a contact offer

- call\_center\_status: is the call center open or closed? This depends mostly on opening hours: some visitors will try to call outside normal opening hours. The call center may also show up as closed if the number of current callers is very high

- device\_type: type of device used by the visitor: mobile/tablet/desktop

- contact: this is the column to predict; it indicates whether the visitor chose to accept the contact offer or not