ONLINE ADVERTISING DASHBOARD

1. INTRODUCTION

As people spend significant amounts of time on the internet, businesses are increasingly turning to *online advertising* to gain new customers and to promote their products and services. To enable this, many popular websites, including search engines, media outlets, blogs and social media platforms, allow companies to place adverts next to their regular content.

Typically, this advertising space is sold using *ad auctions*, which take place whenever a website user views a page. Competing companies submit bids for this particular user via an automated script and the best bids are chosen based on the amount of the bid itself and other relevant factors. Due to the high frequency and speed of these auctions, this mechanism is also known as *real-time bidding*. The final cost to a winning bidder depends on the level of competition and so can vary from one user to the next. The cost for a winning bid may be charged when the ad is shown (known as *pay-per-impression*) or when the user clicks on it (known as *pay-per-click*).

The marketplace for online advertising is complex. Many large websites offer their own auction and bidding mechanisms (such as Google or Facebook), but most smaller websites outsource the selling of their advertising space to *ad networks*, which act as intermediaries between large numbers of websites and advertisers. In order to deal with this complexity, many businesses rely on *online marketing agencies,* which strategically participate in a wide range of ad auctions to maximise their client's return on investment.

2. PROBLEM

Your customer is a small online marketing agency that designs custom advertising campaigns for its clients. The agency would like you to develop a tool that will allow its clients to evaluate the success of advertising campaigns. Specifically, this tool needs to be able to read data gathered during campaigns and then display key metrics about the campaign on a user-friendly and clear dashboard.

Besides providing a quick summary of the overall performance of a specific campaign, such a dashboard should also allow the agency's clients to view performance metrics over time and within particular audience segments or contexts. Such fine-grained information about performance will allow the clients to detect important trends, for example, revealing times of the day when customers tend to respond particularly well to a campaign, identifying sudden changes in performance and highlighting the most engaged audience segments.

3. DEFINITIONS

These are a number of key terms that should be used correctly in your application:

Acquisition: same as *conversion*.

Bounce: A user clicks on an ad, but then fails to interact with the website (typically detected when a user navigates away from the website after a short time, or when only a single page has been viewed).

Bounce Rate: The average number of bounces per click.

Campaign: An effort by the marketing agency to gain exposure for a client's website by participating in a range of ad auctions offered by different providers and networks. Bid amounts, keywords and other variables will be tailored to the client's needs.

Click: A click occurs when a user clicks on an ad that is shown to them.

Click Cost: The cost of a particular click (usually determined through an auction process).

Click-through-rate (CTR): The average number of clicks per impression.

Conversion: A conversion, or acquisition, occurs when a user clicks and then acts on an ad. The specific definition of an action depends on the campaign (e.g., buying a product, registering as a new customer or joining a mailing list).

Conversion Rate: The average number of conversions per click.

Cost-per-acquisition (CPA): The average amount of money spent on an advertising campaign for each acquisition (i.e., conversion).

Cost-per-click (CPC): The average amount of money spent on an advertising campaign for each click.

Cost-per-thousand impressions (CPM): The average amount of money spent on an advertising campaign for every one thousand impressions.

Impression: An impression occurs whenever an ad is shown to a user, regardless of whether they click on it.

Uniques: The number of unique users that click on an ad during the course of a campaign.

3. DATA

For each campaign, the following three log files will be provided:

1) **Impression Log:**
   This comma-separated values (CSV) log file contains information recorded by the agency about the *impressions* generated during the campaign. It contains the following data (where each row corresponds to an impression):

Column	Header	Description
1	Date	Date and time of the impression. Format: 2015-01-20 16:12:47
2	ID	An ID to uniquely identify a particular user.
3	Gender	Gender of user: Male / Female
4	Age	Age group of user: <25 / 25-34 / 35-44 / 45-54 / >54
5	Income	Income of user: Low / Medium / High
6	Context	Context of ad: News / Shopping / Social Media / Blog / Hobbies / Travel

7	Impression Cost	Cost of this impression (in pence).

2) Click Log:

This CSV log file contains information recorded by the agency about the *clicks* generated during the campaign. It contains the following data (where each row corresponds to a click):

Column	Header	Description
1	Date	Date and time of the click on the ad. Format: 2015-01-20 16:12:47
2	ID	Unique ID of user (matches impression log above).
3	Click Cost	Cost of this click (in pence).

3) Server Log:

This CSV log file will be provided by the agency's client about the users' website interactions after clicking on an ad. It contains the following data (where each row corresponds to a user's interactions during one visit).

Column	Header	Description
1	Entry Date	Date and time of arriving on the website. Format: 2015-01-20 16:12:47
2	ID	Unique ID of user (matches logs above).
3	Exit Date	Date and time of navigating away from the website. Format: 2015-01-20 16:29:30 May be "n/a".
4	Pages Viewed	Number of pages user viewed during visit.
5	Conversion	Whether a conversion took place: Yes / No

4. REQUIREMENTS

The mandatory requirements are as follows:

1) The system should be able to read in data relating to a campaign given the three input files described above.

2) The system should be able to compute and display the following key metrics of a campaign:
   a. Number of Impressions
   b. Number of Clicks
   c. Number of Uniques
   d. Number of Bounces
   e. Number of Conversions
   f. Total Cost
   g. CTR
   h. CPA
   i. CPC
   j. CPM
   k. Bounce Rate

3) The system should be able to display charts of these metrics over time. One data point on these charts should correspond to the metric computed over a specific time interval (e.g., one hour, one day or one week).

4) The user should be able to control the time granularity of these time charts.

5) The system should be able to display a histogram of the click costs (showing the distribution of costs).

6) The user should be able to filter the metrics and charts by:
   a. Date range
   b. Audience segments (by gender, age or income)
   c. Context

7) There should be some functionality to directly compare metrics or charts with different filters applied to them (e.g., to compare the CTR within one audience segment to another, or to compare the number of uniques within two distinct time periods). This could be achieved by overlaying graphs or by displaying multiple graphs next to each other.

8) The user should be able to define how a bounce is registered (either depending on the time spent on the website or on the number of pages visited).

9) The software should be fast and responsive. It should be able to handle data sets containing millions of impressions and spanning several months. Once the initial data is loaded, updating the graphs (e.g., by applying filters or switching metrics) should take no more than a few seconds.

Note: the customer would expect you to use a third-party library for displaying charts. Possible options for this include JFreeChart (http://www.jfree.org/jfreechart/) and JavaFX Charts (http://docs.oracle.com/javafx/2/charts/chart-overview.htm).

5. OPTIONAL EXTENSIONS

You are free to implement additional extensions beyond the requirements above. These should be discussed with your supervisor, but could include:

1) Ability to load and compare data from multiple campaigns.

2) Chart displaying the performance metrics per time of day or per day of the week.

3) Facility for saving summary charts to file (as an image or pdf file).

4) Printing functionality.

5) Ability to customise the appearance of the application (e.g., by modifying colour schemes or font sizes).