INSTRUCTIONS

The goal of this exercise is to simulate the creation of a report, by merging two datasets together, and drawing some basic insights. Given the stated assumptions, please provide your answers to the following 4 questions, as well as your code. Use of python, particularly with Pandas, is encouraged.

**Source1.csv:**

* Each “campaign” contains three elements, separated by the delimiter “\_”. The first element represents an initiative, the second represents an audience, and third represents an asset.
  + “A\_B\_C” means the initiative is A, the audience is B, and the asset is C
* Each “actions” value contains a list of dictionaries, where each element has an action and a type. For example {"x": 63, "action": "like"} means that there were 63 likes of type x.

**Source2.csv**

* Each “campaign” contains the same three elements (initiative, audience, asset), separated by the same delimiter “\_”, but in this case the order of the elements is random.

**Assumptions:**

* A “campaign” is a unique combination of Initiative, Asset and Audience
* CPM = spend/impressions\*1000
* CPV = spend/views ONLY for campaigns with an object\_type of video. Ignore spend and views for all other object\_types in calculating CPV.
* All campaigns are represented for each day in source1.csv
* There may be missing or duplicate campaigns in source2.csv
* For all questions, ignore actions that aren’t of type X or Y.

**Questions:**

1. How many unique campaigns ran in February?  **125**
2. What is the total number of conversions on plants?
3. What audience, asset combination had the least expensive conversions?
4. What was the total cost per video view?