Advertisement banners are displayed to users in a mobile application (app\_id) in a country (country code) from an advertiser (advertiser\_id). When this happens, an impression event is recorded and stored. Optionally, if the user clicks on the banner, a click event is recorded. Revenue is generated only in the case of a click being triggered.

**Input**

**Arguments**

Your application should accept **2 lists of file names** with click and impression events.

**Impression event schema**

* id (string): a UUID that identifies the impression.
* app\_id (integer): an identifier of the application showing the impression.
* country\_code (string): a 2-letter code for the country. It doesn't necessarily comply to any standard like ISO 3166.
* advertiser\_id (integer): an identifier of the advertiser that bought the impression.

Example data can be found on impressions.json.

**Click event schema**

* impression\_id (string): a reference to the UUID of the impression where the click was produced.
* revenue (double): the quantity of money paid by the advertiser when the click is tracked.

Example data can be found on clicks.json.

**Goals**

**1. Read events stored in JSON files**

Read and parse the events for both impressions and clicks from the provided JSON files in your entry point. Some events may not comply with the provided schema. You can use the library of your choice to perform the JSON parsing.

**2. Calculate metrics for some dimensions**

The business team wants to check how some metrics perform depending on a few dimensions. For example, they would like to check how applications are performing depending on the country. This will be very useful for them, as they will be able to spot new opportunities or countries that are performing poorly.

Metrics:

* Count of impressions
* Count of clicks
* Sum of revenue

Dimensions:

* app\_id
* country\_code

Please, write the output to a JSON file using the following format:

[

{

"app\_id": 1,

"country\_code": "US",

"impressions": 102,

"clicks": 12,

"revenue": 10.2

},

...

]

**3. Make a recommendation for the top 5 advertiser\_ids to display for each app and country combination.**

Now, the business team wants to know which are the top advertisers for each application and country. This will allow them to focus their effort on this advertisers. To measure performance, we will check for the highest rate of revenue/impressions. That is, the advertisers that, on average, pay more per impression.

Output fields:

* app\_id
* country\_code
* recomended\_advertiser\_ids (list of top 5 advertiser ids with the highest revenue per impression rate in this application and country).

Please, write the output to a JSON file using the following format:

[

{

"app\_id": 1,

"country\_code": "US",

"recommended\_advertiser\_ids": [32, 12, 45, 4, 1]

}

]

**Technical requirements**

* Write your application using the Scala programming language. You can choose the build tool of your choice.
* You can use a library of your choice to parse JSON and program arguments.
* Please, don't use any data processing framework (Spark, Flink, Akka...) for goals 1 and 2. You can use them for the third one.
* Your application will be running on a single instance with 8 cores.

ing the following format:

[

{

"app\_id": 1,

"country\_code": "US",

"impressions": 102,

"clicks": 12,

"revenue": 10.2

},

...

]

3. Make a recommendation for the top 5 advertiser\_ids to display for each app and country combination.

Now, the business team wants to know which are the top advertisers for each application and country. This will allow them to focus their effort on this advertisers. To measure performance, we will check for the highest rate of revenue/impressions. That is, the advertisers that, on average, pay more per impression.

Output fields:

app\_id

country\_code

recomended\_advertiser\_ids (list of top 5 advertiser ids with the highest revenue per impression rate in this application and country).

Please, write the output to a JSON file using the following format:

[

{

"app\_id": 1,

"country\_code": "US",

"recommended\_advertiser\_ids": [32, 12, 45, 4, 1]

}

]

Technical requirements

Write your application using the Scala programming language. You can choose the build tool of your choice.

You can use a library of your choice to parse JSON and program arguments.

Please, don't use any data processing framework (Spark, Flink, Akka...) for goals 1 and 2. You can use them for the third one.

Your application will be running on a single instance with 8 cores.

Clicks.json

[

{

"impression\_id": "97dd2a0f-6d42-4c63-8cd6-5270c19f20d6",

"revenue": 2.091225600111518

},

{

"impression\_id": "43bd7feb-3fea-40b4-a140-d01a35ec1f73",

"revenue": 2.4794577548980876

}

]

Impressions.json

[

{

"app\_id": 32,

"advertiser\_id": 8,

"country\_code": "UK",

"id": "a39747e8-9c58-41db-8f9f-27963bc248b5"

},

{

"app\_id": 30,

"advertiser\_id": 17,

"country\_code": null,

"id": "5deacf2d-833a-4549-a398-20a0abeec0bc"

}

]