

## Assignment Review

In this review, I will describe the overall process, specific task implementations, challenges encountered, and key lessons learned from this project. My approach to the assignment was systematic, dealing with each task in the order they were presented. I began with creating HashMaps for each task, processing the relevant fields from the input file using regex, and then implementing methods to save the results to the corresponding reports. With this approach, I wanted to maintain clarity and organization throughout the project.

The initial task was relatively straightforward. I used a HashMap to store the number of apps per category, incrementing the count by 1 each time a category key appeared in the file. This initial step set up the framework for the following analyses.

Extracting the top 100 companies required parsing App IDs and using regex to identify company names. After incrementing the number of apps per company each time a company key appeared, I then sorted the resulting map entries by values, providing a clear picture of the most productive companies in the Play Store.

To get the top 3 developers, at first, I filtered out developers associated with the company they wrote the app for, by comparing the first two words from the App ID field with the first two words from the Developer Email field (from right to left side). After increasing the number of apps per developer, the map entries are sorted by values, in descending order.

For the app purchase analysis, I created two reports, for \$1000 and \$10000 budgets. The map entries are then sorted by app prices. If the app price is less than or equal to the budget, the price is subtracted from the budget, and the entry key is written into the report.

Distinguishing between free and paid apps for total download counts was achieved using true/false keys in a map. Each time a key appears in the file, the value from the installs field is added to its corresponding map value.

The most challenging part of the assignment was deciding how to start with the project and create its structure.

The part I thought would be very simple is the last task. At first, downloads were of type Integer, and I could not see why some lines were thrown as exceptions. It took me some time to realize that the type should be Long instead, because of the limited range of the Integer type. I recognized the need to pay attention to details and the importance of considering data types while doing this task. Additionally, identifying the top 3 developers was more complex than expected.

If I was starting the assignment again, I would probably pay more attention to the details, like making the code look cleaner and easier to read. Optimization is also something I should improve.

Ajša Herceglija