

Problem statement:

A DoorDash Drive merchant wants us to run a marketing campaign targeting some of their users and has an agreement where they plan to send us all of their customer phone numbers across the globe that should be a part of this campaign. Unfortunately, the merchant's data isn't very clean and multiple phone numbers are defined as one string! The merchant is also not willing to clean up their data, so we need to work with it as-is.

For example, the string "(Home) 415-415-4155 (Cell) 415-514-5145" would represent two different phone numbers with different designations (Home vs. Cell). As the backend engineer in charge of this project, you'll need to build an API for the merchant to use.

This API is expected to parse an input string (format: "(Home) 415-415-4155 (Cell) 415-514-5145") to generate some data structure containing each phone number type and phone number, and persist this data structure to some data store.

Requirements:

- a. Data that gets saved as a part of this API should be persisted across application restarts.
- b. You cannot use regular expressions or third party libraries to parse the string.
- c. You need to persist how many times each distinct phone-type pair has been encountered before.
- d. This API should return a list of pairs of each phone number + type

Example input 1: POST to /phone-numbers with JSON data {"raw_phone_numbers": "(Home) 415-415-4155 (Cell) 415-123-4567"}

Example output 1: {"results": [
 {"id": "someid", "phone_number": "4154154155", "phone_type": "home", "occurrences": 1},
 {"id": "someotherid", "phone_number": "4151234567", "phone_type": "cell", "occurrences": 1}
]} // The occurrences for each result should be **1** since this is the first request containing those phone number + type pairs.

Example input 2: POST to /phone-numbers with JSON data {"raw_phone_numbers": "(Home) 415-415-4155 (Cell) 415-123-4567"}

Example output 2: {"results": [
 {"id": "someid", "phone_number": "4154154155", "phone_type": "home", "occurrences": 2},
 {"id": "someotherid", "phone_number": "4151234567", "phone_type": "cell", "occurrences": 2}
]} // Note that occurrences = **2** because this is the 2nd time each number + type has been encountered.

Implementation guidance:

1. To begin work on this problem, you should create a private github repository (free), a branch called `interview` to push changes to, and invite your.interviewer@doordash.com as a collaborator. You are welcome to use our project template's code (Kotlin) as a starting point, but otherwise feel free to create your own project in whatever language/framework you are most comfortable with.

2. For any clarifying questions about the problem statement, feel free to reach out to your interviewer over email! They're here to help.
3. To submit your work when you're done, send your interviewer a link (via email) to a pull request from the `interview` branch so that they can collaborate with you on a code review.
4. Each API should return a result in no more than 100ms.
5. These APIs do not need to be deployed anywhere, they just need to run locally.