

# CategoriesCount ReadMe

**Introduction:** This part counts the number of businesses in each category in Phoenix.

**Input:** yelp\_academic\_dataset\_business.json

**Mapper:** Since the input data is in JSON format. I used JsonParsing.java which uses gson-2.8.0.jar to parse the JSON input to hash map.

Then I can use the key of hash map to get the category of each business record. Because in the business dataset, column category is an array, thus I take every category out and check if the business is in Phoenix. If so, write it as the output of the Mapper Class as <key: category name, value: 1>

**Reducer:** The reducer class simply takes the output of Mapper class and sum up the value for the same key(category).

**Output:** CatCountResult\_Phoenix

**Screenshot:** Category\_ScreenShot