

# ContentSquare Data Engineering Test

## Auto-complete

You are given a list of keywords (on the next page). Write a program that will offer up to 4 suggested keywords from the list, starting with the letters typed, **not case sensitive**. Similar to Google Autocomplete, except that results must be in alphabetical order. You will find examples at the end of this document.

## Requirements

- You can use any programming language of your choice, preferably Scala or Java
- Do not make a web interface for input and output, a simple method taking a string and returning a list of strings is fine
- Do not use databases and complex libraries, write the completion logic yourself
- Think about optimization and algorithmic complexity, explaining the tradeoffs that you chose to made, if any

## Deliverable

- Sources of your program, commented when necessary
- Other artifacts and files required to run the program (like the list of keywords, if your program loads it)

## Additional Questions

No code expected, please give your answers in a Readme file

- What would you change if the list of keywords was much larger (300 Gb) ? Please explain and describe the concepts that would allow to handle this if you decide to use specific tools (frameworks, databases...)
- What would you change if the requirements were to match any portion of the keywords (for example, given the string “pro”, the program could suggest the keyword “reprobe”) ?

# List of keywords

project runway  
pinterest  
river  
kayak  
progenex  
progeria  
pg&e  
project free tv  
bank  
proactive  
progesterone  
press democrat  
priceline  
pandora  
reprobe  
paypal

## Examples

Autocomplete(**"p"**)

pandora  
paypal  
pg&e  
pinterest

---

Autocomplete(**"pr"**)

press democrat  
priceline  
proactive  
progenex

---

Autocomplete(**“pro”**)

proactive  
progenex  
progeria  
progesterone

---

Autocomplete(**“prog”**)

progenex  
progeria  
progesterone