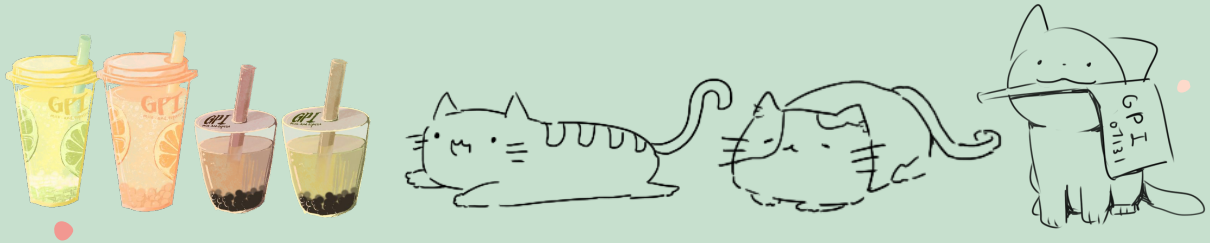
Lec 8 : ~~Boba~~ Regex and ~~Cat~~ Grep



F21 07131 GPI Lecture

Announcements

- Resume Review extratation this weekend! (3rd to last one 👀)
  Come in with a resume!!
- Collab part 2 grade should be out now!

Pet tax!

Globfield

Pregex

Igreprus

What we want to achieve	Glob/Range
Make a file for all 30 houses numbered 1 ~ 30	$ touch house{1..30}
Throw eggs to the houses 7131 to 7139, if it exists	$ ./throwEgg house713?
Send every person in scs 1x free boba voucher	$ ./sendFreeBobaVoucher *.scsPerson
... except Tom and Veronica	???? (impossible to do efficiently)
Vouchers for Tom and Veronica should allow for unlimited boba! We want something that will accept/match “”, “boba”, “bobaboba”, “bobabobaboba”... etc (boba only for now... not enough budget)	???? (impossible)

Glob's kinda useless ... Do we have something better?

Introducing ... REGEX!

- “Regular Expression”
  - Patterns that match against certain strings
  - Different from globs!
  - Compatible with many applications
- But why are they called regular expressions?
  - For interesting theoretical reasons (that you will learn later)
- Example (that we’ll dissect later)
  - (\d{3}-?){2}\d{4}

Example - boba shop phone numbers

- Multiple possible strings
  - 123-456-7890
  - 1234567890
  - 456-789-1234
- But the formats follow a few patterns
  - ###-###-####
  - ##########

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu vehicula leo felis nec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apien. Fusce quam diam, imperdiet nec sollicitudin a, egestas ut ligula2093523. Etiam h3290852e329505ndrerit nunc et eros 31905-031831ultrices eleifen23985-69d.
Sed diam neque, ullamcorper sit amet dapib390-193us ac,410-849-3958 porta a enim. Aliquam semper dictum nunc, et faucibus ante posuere ut. Vivam32890388us a sapien commod31531o, sagittis velit eget, venenae9-=8528-235tis ligula. Pellente30953sque eget massa in nisi semper aliquam. Praesen32946t 3902135in facilisis mauris, vel mattis urna. Aenean finibus venenatis neque fringilla facilisis. Du4098620-2is sed metus finibuse90835-18, eff496=29icitur neque vitae, suscipit lectus. Duis semper sapien urna, quis rhoncus justo dapib31901390us pulvinar. 13436Nunc posuere tortor qui3903-19193s nisi bibendum, sed maximus mauris sodales. Phasellus vitae398-193939 metus quis diam laoreet euismoec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sad. Vivamus tempor vehicula bibeec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a392 dictum enim. Nulla volutpat sit amet risus rutrum varius. Pellentesque egestas a tellus eget egestas. Suspendisse ut ipsum sodales eros cursus pretium vel ut nisl. Pel3141577lentesque laoreet aliquet faucibus.
Morbi ultricies diam in commodo aliquet. Curabitur ac aliquet leo. Sed eros mi, vulputate bibendum ligula vel, sollicitudin 193-948-3195 congue quam. Cras et bibendum neque. Duis euismod luctus dolor, at facilisis risus39393930-1 aliquet ut. Fusce quis ipsum matt3939is, molestie ex a, tincidunt velit. Nulla iD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienn orci dapibus, congue enim sit amet, bibendum ante. Praesent facilisis massa eget cursus accumsan.
enenatis elit. Nunc elit lectus, tempus eu gravida id, elementum at nisl. Etiam vitae nulla nec nibh egestas sodales ut sit amet dui. Pellentesque id tortor mi.
Lorem i-psum dolor sit amet, consectetur adipiscing elit. 3949500395Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu- vehicula leo D32908153uis ut enim porttitor, a3-9201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu1309385is, aliquam hendrerit s19035-6913-35299043092apienfelis nec enim. Nulla fringilla enim et ipsum commodo tinc--idunt. Ut sed ullamcorper nisi. Duis ut enim porttitor, aliquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna quis, aliquam hendrerit sapien. Fusce quaD32908153uis ut enim porttitor, a39201-6932liquet sapi-en rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienm diam, imperdiet nec sollicitudin a, egestas ut ligula. Etiam hendrerit nunc et eros ultrices eleifend.-In nec dignissim arcu. Donec sed blandit tortor, sed vD32908153uis ut enim porttitor, a39201-693-2liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35Morbi ultricies diam in commodo aliquet. Curabitur ac alD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913--352990-4309--2apienD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s1-9035-6913-352990-43092apieniquet leo. Sed eros mi, vulp-utate bibendum ligula vel, sollicitudin congue quam. Cras et bibendum neque. Duis euismod luctus dolor, at facilisis risus aliquet- ut. Fusce quis ipsum mattis-, molestie ex a, tincidunt velit. Nul-la in orci dapi-bus, congue enim sit amet, bibendum ante. Praesent facilisis massa eget cursus accumsan.
Mauris condimentum, nisi vel varius posuere, metus ante aliquam nibh, sed pharetra tellus ex vel enim. Phasellus viverra hend-rerit 113-295-1935ligula ac feugiat. Suspendisse ut mattis lacus. Curabitur nec condimentum justD32908153uis ut enim porttitor, a39201-6932liquet sa-pien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apieno, vel convallis diam. Mauris tristique venenatis elit. Nunc elit lectus, tempus eu gravida id, elementum at nisl. Etiam vitae nulla nec nibh egestas sodales ut sit amet dui. Pellentesque id tortor mi. r. 13436Nunc posuere tortor qui3903-19193s nisi bibendum, sed maximus mauris sodales. Phasellus vitae398-193939 metus quis diam laoreet euismoec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sad. Vivamus tempor vehicula bibeec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sandum. VestibLorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu vehicula leo felis nec enim. Nulla fringilla enim et ipsum commodo tincidunt. Ut sed ullamcoD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienrper n-isi. Duis ut enim porttitor, aliquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna quis, aliquam hendrerit sapien. Fusce quam diam, imperdiet nec sollicitudin a, egestas ut ligula. Etiam hendrerit nunc et eros ultrices eleifend.
Sed diam neque, ullamcorper sit amet dapibus ac, p4861930135orta a enim. Aliquam semper dictum nu-nc, et fauci-bus ante posuere ut. Vivamus a sapien commodo, sagittis velit eget, venenatis ligula. Pellentesque eget massa in nisi semper aliquam. Praesent in facilisis mauris, vel mattis urna. Aenean f-inibus venenatis neque fringilla facilisis. Duis sed metus finibus, efficitur neque vitae, suscipit lectus. Duis semper sapien urna, quis rhoncus justo dapibus pulvinar. Nunc posuere tortor quis nisi bibendum, sed maximus mauris sodales. Phasellus vitae metus quis diam laoreet euismod. Vivamus tempor vehicula bibendum. Vestibulum aliquam nulla eu varius bibendum. Nulla suscipit nec felis quis finibus.Sed diam neque, ullamcorper sit amet dapibus ac, porta a enim. Aliquam semper dictum nunc, et faucibus ante posuere ut. Vivamus a sapien commodo, sagittis velit eget, venenatis li-gula. Pellentesque eget massa in nisi semper aliquamD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apien. Praesent in facilisD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienis mauris, vel mattis urna-. Aenean finibus venenatis neque fringilla facilisis. Duis sed metus finibus, efficitur neque vit-ae, suscipit lectus. Duis semper sapien urn--a, quis rhoncus justo dapibus pulvinar. Nunc posuere tortor quis nisi bibendum, sed maximus mauris sodales. Phasellus vitae metus quis diam laoreet euismod. Vivamus tempor vehicula bibendum. Ves-tibulum aliquam nulla eu varius bibendum. Nulla suscipit nec felis quis finibus.--In nec dignissim arcu. Donec sed blandit tortor, sed varius orci. Sed euismod accumsan erat. Donec risus massa, tristique n3859393964on euismod id, faucibus quis neque. Morbi magna enim, porttitor vel leo sit amet, interdum lacinia purus. Vivamus lobortis, nunc eu molestie sodales, felis nisl consectetur leo, efficitur laoreet risus dui vel nulla. Duis tempor nec est id egestas. Nulla ac dictum enim. Nulla volutpat sit amet risus rutrum varius. Pellentesque egestas a tellus eget egestas. Suspendisse ut ipsum sodales eros cursus pretium vel ut nisl. Pellentesque laoreet aliquet faucibus.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu vehicula leo felis nec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apien. Fusce quam diam, imperdiet nec sollicitudin a, egestas ut ligula2093523. Etiam h3290852e329505ndrerit nunc et eros 31905-031831ultrices eleifen23985-69d.
Sed diam neque, ullamcorper sit amet dapib390-193us ac, **410-849-3958** porta a enim. Aliquam semper dictum nunc, et faucibus ante posuere ut. Vivam32890388us a sapien commod31531o, sagittis velit eget, venenae9-=8528-235tis ligula. Pellente30953sque eget massa in nisi semper aliquam. Praesen32946t 3902135in facilisis mauris, vel mattis urna. Aenean finibus venenatis neque fringilla facilisis. Du4098620-2is sed metus finibuse90835-18, eff496=29icitur neque vitae, suscipit lectus. Duis semper sapien urna, quis rhoncus justo dapib31901390us pulvinar. 13436Nunc posuere tortor qui3903-19193s nisi bibendum, sed maximus mauris sodales. Phasellus vitae398-193939 metus quis diam laoreet euismoec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sad. Vivamus tempor vehicula bibeec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a392 dictum enim. Nulla volutpat sit amet risus rutrum varius. Pellentesque egestas a tellus eget egestas. Suspendisse ut ipsum sodales eros cursus pretium vel ut nisl. Pel3141577lentesque laoreet aliquet faucibus.
Morbi ultricies diam in commodo aliquet. Curabitur ac aliquet leo. Sed eros mi, vulputate bibendum ligula vel, sollicitudin **193-948-3195** congue quam. Cras et bibendum neque. Duis euismod luctus dolor, at facilisis risus39393930-1 aliquet ut. Fusce quis ipsum matt3939is, molestie ex a, tincidunt velit. Nulla iD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienn orci dapibus, congue enim sit amet, bibendum ante. Praesent facilisis massa eget cursus accumsan.
enenatis elit. Nunc elit lectus, tempus eu gravida id, elementum at nisl. Etiam vitae nulla nec nibh egestas sodales ut sit amet dui. Pellentesque id tortor mi.
Lorem i-psum dolor sit amet, consectetur adipiscing elit. **3949500395**Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu- vehicula leo D32908153uis ut enim porttitor, a3-9201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu1309385is, aliquam hendrerit s19035-6913-35299043092apienfelis nec enim. Nulla fringilla enim et ipsum commodo tinc--idunt. Ut sed ullamcorper nisi. Duis ut enim porttitor, aliquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna quis, aliquam hendrerit sapien. Fusce quaD32908153uis ut enim porttitor, a39201-6932liquet sapi-en rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienm diam, imperdiet nec sollicitudin a, egestas ut ligula. Etiam hendrerit nunc et eros ultrices eleifend.-In nec dignissim arcu. Donec sed blandit tortor, sed vD32908153uis ut enim porttitor, a39201-693-2liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35Morbi ultricies diam in commodo aliquet. Curabitur ac alD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913--352990-4309--2apienD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s1-9035-6913-352990-43092apieniquet leo. Sed eros mi, vulp-utate bibendum ligula vel, sollicitudin congue quam. Cras et bibendum neque. Duis euismod luctus dolor, at facilisis risus aliquet- ut. Fusce quis ipsum mattis-, molestie ex a, tincidunt velit. Nul-la in orci dapi-bus, congue enim sit amet, bibendum ante. Praesent facilisis massa eget cursus accumsan.
Mauris condimentum, nisi vel varius posuere, metus ante aliquam nibh, sed pharetra tellus ex vel enim. Phasellus viverra hend-rerit **113-295-1935**ligula ac feugiat. Suspendisse ut mattis lacus. Curabitur nec condimentum justD32908153uis ut enim porttitor, a39201-6932liquet sa-pien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apieno, vel convallis diam. Mauris tristique venenatis elit. Nunc elit lectus, tempus eu gravida id, elementum at nisl. Etiam vitae nulla nec nibh egestas sodales ut sit amet dui. Pellentesque id tortor mi. r. 13436Nunc posuere tortor qui3903-19193s nisi bibendum, sed maximus mauris sodales. Phasellus vitae398-193939 metus quis diam laoreet euismoec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sad. Vivamus tempor vehicula bibeec enim. N29305ulla frin395082gilla enim et ipsum c32905-393ommodo tincidunt. Ut sed ullamcorper nisi. D32908153uis ut enim porttitor, a39201-6932liquet sandum. VestibLorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam elementum interdum ligula eu tempor. Etiam feugiat, lectus a hendrerit vehicula, lacus tellus gravida justo, eu vehicula leo felis nec enim. Nulla fringilla enim et ipsum commodo tincidunt. Ut sed ullamcoD32908153uis ut enim porttitor, a39201-6932liquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apienrper n-isi. Duis ut enim porttitor, aliquet sapien rhoncus, porttitor elit. Sed ex nunc, porta a magna qu aliquam hendrerit sapien. Fusce quam diam, imperdiet nec sollicitudin a, egestas ut ligula. Etiam hendrerit nunc et eros ultrices eleifend.
Sed diam neque, ullamcorper sit amet dapibus ac, p**4861930135**orta a enim. Aliquam semper di nu-nc, et fauci-bus a posuere ut. V a sapien commodo, sagittis velit eget, venenatis ligula. Pellentesque eget massa in nisi semper aliquam. Praesent in facilisis ma mattis urna. A inibus neque fringilla facilisis. Duis sed metus finibus, efficitur neque vitae, suscipit lectus. Duis semper sapien urna, quis rhoncus justo dapi maximus mauris sodales. Phasellus vitae metus quis diam laoreet euismod. Vivamus tempor vehicula bibendum. Ves uscipit nec felis quis finibus.Sed diam neque, ullamcorper sit amet dapibus ac, porta a enim. Aliquam semper dictum nunc, et sagittis velit eget, venenatis li-gula. Pellentesque eget massa in nisi semper aliquamD32908153uis unc, porta a magna qu130985is, aliquam hendrerit s19035-6913-35299043092apien. Praesent in fac honcus, porttitor elit. Sed ex nunc, porta a magna qu130985is, aliquam hendrerit s19035-6913-3529904309 eque fringilla facilisis. Duis sed metus finibus, efficitur neque vit-ae, suscipit lectus. Duis semper sapien urn--a isi bibendum, sed maximus mauris sodales. Phasellus vitae metus quis diam laoreet euismod. Vivamus varius bibendum. Nulla suscipit nec felis quis finibus.--In nec dignissim arcu. Donec sed blandit tortor varius orci. Se , tristique n**3859393964**on euismod id, faucibus quis neque. Morbi magna enim, porttitor vel leo sit amet, interdum lacinia mor ales, felis nisl consectetur leo, efficitur laoreet risus dui vel nulla. Duis tempor nec est id egestas. Nulla ac dictum enim. Nulla us ru us. Pellen e egestas a tellus eget egestas. Suspendisse ut ipsum sodales eros cursus pretium vel ut nisl. Pellentesque laoreet aliquet faucibus.

(\d{3}-?){2}\d{4}

Introducing ... REGEX!

- Normal characters
- Character classes
- Quantifiers
- Groups
- Other special expressions

Parts of a regular expression

- Normal characters
  - boba – matches "boba" 
- Character classes
  - [sml] – matches "s", "m", or "l" (no, not standard meta language, it's small, medium, or large)
  - \d – matches a digit
  - . – matches any character

Character Classes

“Boba is my character.
Boba is my personality”

Class	Matches
[abc]	a or b or c
[a-z]	Any single lowercase letter
\s	whitespace
\d	digit
\w	alphanumeric char
.	any single character (ie, wildcard)

More Character Classes

Class	Matches
[abc]	a or b or c
[a-z]	a lowercase letter
\s	whitespace
\d	digit
\w	alphanumeric char
.	any single character (ie, wildcard)

Class	Matches
[^abc]	Not any of a, b or c
[^A-Za-z]	Not a character
\S	Not a whitespace
\D	Any non-digit
\W	Any non-alphanumeric char
\.	Just a period

Parts of a regular expression

- Up until now, we've seen character classes that only match to 1 single character. How do we specify ~*quantity*~?

More Quantifiers

"How many cups of boba do you want?"
"yes"

Quantifier	Matches
a?	Zero or one (ie. optional)
a*	Zero or more
a+	One or more
a{3}	Exactly 3
a{3,}	3 or more
a{3,6}	Between 3 and 6 (inclusive)

Parts of a regular expression

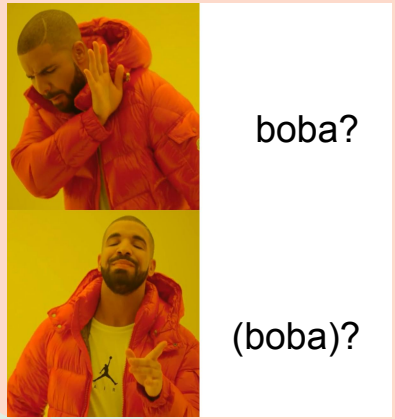
- Up until now, we've seen character classes that only match to 1 single character. How do we specify ~*quantity*~?

- Quantifiers act on the preceding character...
  - boba? - matches "bob" or "boba"
  - boba* – matches "bob", "boba", " bobaaaaa", etc.

or

Parts of a regular expression

- So far, the quantifiers only acts on the “character” immediately before it.
- Use parentheses for grouping
  - boba* - matches “bob”, “boba”, “bobaaa”...
  - (boba)* - matches “”, “boba”, “bobaboba”, “bobabobaboba” …
  - boba? - matches “bob” or “boba”
  - (boba)? - matches exactly “” or “boba”



Note that pattern matching does NOT have to start from the beginning!

- Pattern: \d (will match any digit)
- Matches: (almost like a search-and-find)
  - abc**1**xyz
  - **1**abcxyz
  - abcxyz**1**

... but you might get false positives

- Pattern: free boba
- Matches:
  - “**free boba** with the purchase of 10 bobas”
  - “**free boba** when you spend $1000 or more”
  - “Breaking news: care**free boba**-lover spends $1000”

Other special expressions

- ^ - Start of string/line (not to be confused with ^ in [^abc])
  - **^(free)** matches only “free” at beginning of the string/line
- $ - End of string/line
  - **(boba)$** matches only “boba” at the end of the string/line

^ - jump to the first non-blank character of the line
$ - jump to the end of the line

Other special expressions

- **(this|that)**
  - Don’t put spaces unless you want to match a space
  - Put as many as you want: **(this|that|here|there)**
  - Can nest or use other character classes! Ex. you can express the above as **(th(is|at)|t?here)**

Class	Matches
[abc]	a or b or c
[a-z]	a lowercase letter
\s	whitespace
\d	digit
\w	alphanumeric char
.	any single character
[^abc]	Not any of a, b or c
[^A-Za-z]	Not a character
\S	Not a whitespace
\D	Any non-digit
\W	Any non-alphanumeric char
\.	Just a period

Quantifier	Matches
a?	Zero or one (ie. optional)
a*	Zero or more
a+	One or more
a{3}	Exactly 3
a{3,}	3 or more
a{3,6}	Between 3 and 6 (inclusive)

Other expressions	Matches
^(start)	“start” at the beginning of line/string
(end)$	“end” at the end of the line/string
(this|that)	“this” or “that”

Example - boba shop phone numbers

- Multiple possible strings
  - 123-456-7890
  - 1234567890
  - 456-789-1234
- But the formats follow a few patterns
  - ###-###-####
  - ##########

Example - boba shop phone numbers

```
(\d{3}-?){2}\d{4}
```

Example - boba shop phone numbers

```
(\d{3}-?){2}\d{4}
```

Matches any 3 digits

Example - boba shop phone numbers

```
(\d{3}-?){2}\d{4}
```

Matches an optional hyphen

Example - boba shop phone numbers

Ex:
123-456-
123456-
123456

```
(\d{3}-?){2}\d{4}
```

Matches 2 groups of 3 digits

Example - boba shop phone numbers

```
(\d{3}-?){2}\d{4}
```

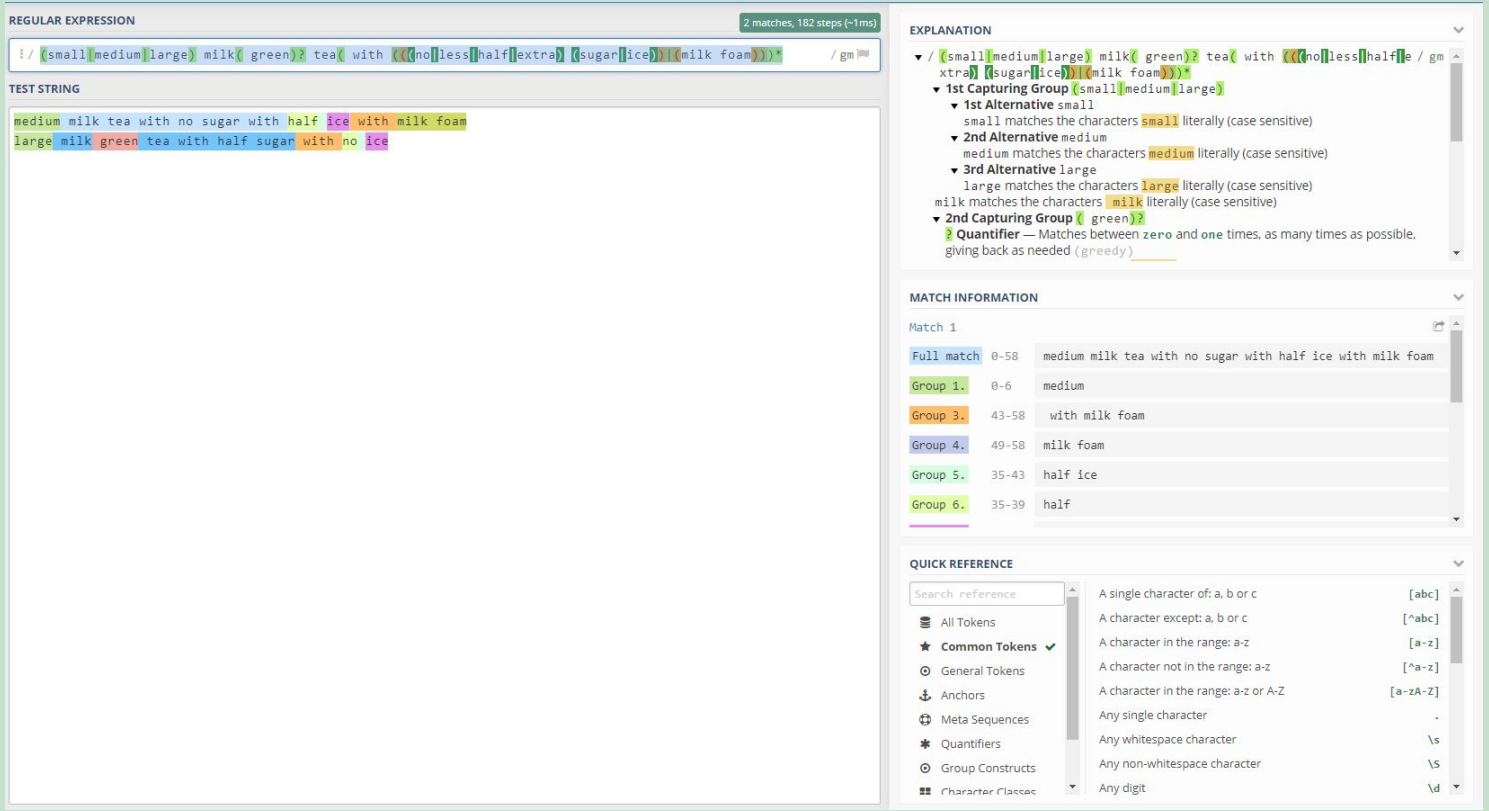
Matches 2 groups of 3 digits, then
4 more digits

Example

```
(small|medium|large) milk( green)? tea(
with (((no|less|half|extra)
(sugar|ice))|(milk foam)))*
```

Match your own favorite drink order!

https://regex101.com/



Quiz!

Matches	Regex
bobabobaboba or **bobaboba**	bobaboba(boba)? or (boba){2,3}
boba any number of times	(boba)*
[any letter][any number] ex: A4	[A-Za-z]\d
example.com website.com etc.	[a-z]+\.com

It's ~~confusion~~ time

Alright let's put glob back into our brain...

Regex vs Globs and ranges

Regex	Glob/Range equivalent
`.`	`?`
`IWant[1-7]CupsOfBoba\.please`	`IWant{1..7}CupsOfBoba.please`
`[sml]`	`{s,m,l}`
`(this|that)`	`{this,that}`
`.*`	`*`
`(ab)*`	`Not possible`

When to use what?

Regex

- Grep, sed, vim
- Searching with regex is also supported in VSCode, Google Sheets, etc
- Useful for parsing or validating data like an email, phone number, password, credit card num etc.

Glob

- In terminal command line, used by shells for matching file and directory names using wildcards. The capabilities of globbing depend on the shell.

GREP

- Bash terminal command!
- **G**lobal **R**egular **E**xpression **P**rint
- Usage: $ grep [flags] ‘pattern’ [file...]
- **$ grep 'needle' haystack.txt**
  - Searches haystack.txt for "needle"s.
- **$ grep -r 'boba' path/to/directory**
  - Searches recursively for "boba"s.
- grep is fast for theoretical reasons you will learn in the future

GREP - note!

- If you have quotes in your pattern, make sure to either escape OR put the whole pattern in another quotation
    - **$ grep '"' file or $ grep \" file**
    - **$ grep "'" file or $ grep \' file**

- If you have spaces in your pattern, put the whole pattern in quotes!
    - **$ grep 'free boba voucher' file**
- Remember that in general, put quotes around your shell command arguments if it has quotes!

most ambitious crossover in history?

use regex for the pattern...

… and globs for the filename(s).

$ grep [flags] ‘pattern’ file

wrap the pattern with quotes!!

(a glob of kittens)

flags

http://linuxcommand.org/lc3_man_pages/grep1.html

-l to print filenames

-i (ignores case when searching)

-c (counts up the number lines that contains a match)

-v (inverse; print out lines that do NOT match)

-n (also print out the line numbers)

-E (use extended regex)

Matching multiple patterns/files

(escaped) pipe to match patterns

Bracket notation to match file names

$ grep “pattern1\|pattern2\|pattern3” {file1, file2, file3}

equivalent to

$ grep -E “pattern1|pattern2|pattern3” {file1, file2, file3}
(-E flag if you don’t want to use ‘\’)

GREP - note!

- There are a lot of flags and edge cases and when-to-use-what (basic vs extended vs perl), so when stuff don't work, try googling it first!
- Example:
    - **$ grep (this|that) file** doesn't work …
    - Need to use egrep (or -E flag): **$ grep -E (this|that) file**
    - Same issue when using braced quantifiers like {3}
- Example:
    - **$ grep \d file** to look for a digit doesn't work… What should I do?
    - Use **$grep [[:digit:]] file** or **$grep -P '\d' file**

SED - stream editor

- Can perform find and replace on a file
- **$ sed 's/find/replace/g' path/to/file**
    - Prints result of replacement to the command line, leaving input untouched
    - You can use regex in find! ex/ **$ sed ‘s/[0-9]//g’ file.txt**

- **$ sed -i 's/find/replace/g' path/to/file**
    - -i for "In place"
    - Edits the file

Globfield

Pregex

Igreprus


cat
hair


CRAFTING WITH CAT HAIR
Cute Handicrafts to Make with Your Cat
by Kaori Tsutaya
translated from the Japanese
by Amy Hirschman

we want to make cat hair crafts

- we want to ~~grep~~grab it quickly (in one go) so we can start crafting
- All cat hair is labeled with the cat's name
  - **$ grep "Globfield" catsBeanBag**
  - **$ grep "Pregex" catsBeanBag**
- We want to get all the cat hair at once
  - **$ grep "Globfield\|Pregex\|Igreprus" catsBeanBag**
  - **$ grep -E "Globfield|Pregex|Igreprus" catsBeanBag**

globfield's cousin?

we want to make cat hair crafts

- There's a lot more hair throughout the apartment...
- To split the work of collecting cat hair, and we need a list of all the rooms(files) that have cat hair and we'll each take an equal amount of rooms.
    - **$ grep -lE "Globfield|Pregex|Igreprus" apartment/***
- Jeremy's responsible for the living room and bedroom! He wants to know exactly where each line each tuft of hair is located.
    - **$ grep -nE "Globfield|Pregex|Igreprus" apartment/{livingRoom,bedRoom}**
- Note how we use globs for directory!

l to **list** all the files
n for line **numbers**

we want to make cat hair crafts

- Friends are coming over, and they might not like to sit on cat hair
  - **$ grep -vE “Globfield|Pregex|Igreprus” apartment/livingRoom**
- Turns out they are allergic, so we want to know which room(file)(s) don’t have cat hair
  - hmm.. let’s try $ grep -vl “Globfield\|Pregex\|Igreprus” apartment/* … not quite
  - http://linuxcommand.org/lc3_man_pages/grep1.html
  - **$ grep -L “Globfield\|Pregex\|Igreprus” apartment/***

l to **list** all the files
v for the in**verse**

Remember...

- Give us feedback!
  http://tinyurl.com/f21-gpi-feedback
- Extratation for this weekend!
  Topic: Resume review. Come in w/ a resume!
- Next week's lecture... Advice & QA

2 labs for 2 weeks!

ZombieLab

- Be careful with escaping correctly
- Don't forget to do **$ chmod +x script.sh** and add **#!/bin/bash**

Hauntlab

- **Remember to put quotes around regex patterns!**
- Experiment with running some commands you think will work at the command line first. Then, when you think you've got an answer that will work, transfer it to your script.