Regular Expression Cheat Sheet (Inspired by Jurafsky Slides)

- Disjunctions

RE	Match	Example Patterns Matched
`[mM]oney`	Money or money	"Money"
`[abc]`	'a', 'b', *or* 'c'	"Investing in Iran"
		"is dangerous business"
`[1234567890]`	any digit	"sitting on $7.5 billion dollars"
		"2005 and 2006, more than "
		"$150 million dollars"
`[\.]`	A period	" 'Run!', he screamed."

- Ranges

RE	Match	Example Patterns Matched
`[A-Z]`	an upper case letter	"Rep. Anthony Weiner
		(D-Brooklyn & Queens)"
`[a-z]`	a lower case letter	"ACORN's"
`[0-9]`	a single digit	"(9th CD) "

- Negations

RE	Match	Example Patterns Matched
`[^A-Z]`	not an upper case letter	"ACORN**'s**"
`[^Ss]`	neither 'S' nor 's'	"**ACORN**'s"
`[^\.]`	not a period	" 'Run!', he screamed."

- Optional Characters: `?`, `*`, `+`

RE	Match	Example Patterns Matched
`colou?r`	Words with `u` 0 or 1 times	"color" or
		"colour "
`oo*h!`	Words with `o` 0 or more times	"oh!" or
		"ooh!" or
		"oooh!"
`o+h!`	Words with `o` 1 or more times	"oh!" or
		"ooh!" or
		"oooooh!" or

- Wild Cards `.`

RE	Match	Example Patterns Matched
`beg.n`	Any word with "beg" then "n"	"begin" or
		"began" or
		"begun" or
		"beggn" (Poor grammar!)

- Start of the line anchor ^, end of the line anchor **$**

RE	**Match**	**Example Patterns Matched**
`^[A-Z]`	Upper case start of line	"Palo Alto" "the town of **P**alo Alto"
`^[^A-Z]`	Not upper case start of line	"the town of Palo Alto" "**P**alo Alto"
`^.`	Start of line	"Palo Alto" "the town of Palo Alto"
`.$`	Identify character that ends a line	"Wait**!**" "This is the end**.**"

- "Or"| statements, Useful short hand

RE	**Match**	**Example Patterns Matched**
yours|mine	Matches"yours" or "mine"	"it's either yours or mine"
\ d	Any digit	"1-Mississippi"
\ D	Any non-digit	"1-Mississippi"
\ s	Any whitespace character	"1, 2"
\ S	Any non-whitespace character	"1, 2"
\ w	Any alpha-numeric	"1-Mississippi "
\ W	Any non-alpha numeric	"1-Mississippi"