REGULAR EXPRESSIONS: REGEXES IN PYTHON

What you will learn:

1. Using regular expressions: when and why?
2. Matching characters and ranges of characters.
3. Controlling the placement of a match.
4. Grouping and repeating in patterns.
5. Using the Python `re` module.
6. Modifying regex searches with flags.

VERSIONS

Note:

- Code samples were tested using: Python 3.8.2
- Subtle change in re.sub() in Python 3.7 where matches are zero-length
- Python3 defaults to Unicode

LANGUAGE WITHIN A LANGUAGE

- A Regular expression, or **regex** for short is:
  - A sequence of characters that define a search pattern
- Used for find and replace
- Concept first formalized by Stephen Cole Kleene in the 1950s
- Variety of syntaxes:
  - POSIX standard
  - Perl syntax, or **PCRE** (Perl Compatible Regular Expression)

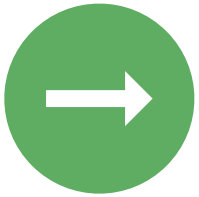
WHAT ARE REGEXES USED FOR?

- Text matching
- Verifying input
- Search and replace
- More powerful versions of `str.split()` and `str.replace()`
- Matching multiple parts of a string and parsing at the same time

SUPPORT

- Used in a variety of UNIX utilities:
  - `sed`, `awk`, `grep`
- Many word processors and text editors
- POSIX standard has two parts:
  - BRE: Basic Regular Expressions
  - ERE: Extended Regular Expressions
- Most major languages:
  - Perl, Java, Javascript, Julia, Python, Ruby, Qt, .NET, PHP, XML Schema
  - Most languages use PCRE with some minor differences

NEXT UP...



Matching characters and ranges

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. Meta-characters
    c. Anchors
    d. Quantifiers
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REGEXES RANGE IN COMPLEXITY

- A plain string match:

```
r"thing"
```

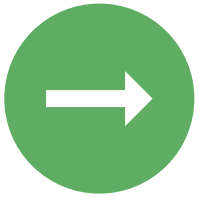
- The RFC822 e-mail address header matching standard:

```
r"(?:(?:\r\n)?[ \t])*(?:(?:(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\]
\000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*))*@(?:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\
["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?
[ \t])*))*|(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+
(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\
\]|\\.)*\](?:(?:\r\n)?[ \t])*))*(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:
[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\
["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\
\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])*(?:(?:(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?
[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?
[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|
\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*|(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?
[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\
[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?
[ \t])*)?(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*\>(?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?
[ \t])*)?(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:
(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:
[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*\>(?:(?:\r\n)?[ \t])*)(?:,
\s*(?:(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:
\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^
[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*|(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\
["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?
[ \t])*))*|(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\
[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*\>(?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\
[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*\>(?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\
[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-
\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*))*\>(?:(?:\r\n)?[ \t])*)*;\s*)"
```

GETTING STARTED

- Learn the regex language before the Python modules
- Python variant
- First lesson:
  - Plain text matching
  - Character classes: choosing from a range of characters

NEXT UP...



Using meta-characters

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. Anchors
    d. Quantifiers
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REGEX LANGUAGE SO FAR...

- A plain string match:

RegEx	Text With Match
`thing`	I put the thing in the place
`spam`	Well, there's spam egg sausage and spam, that's not got much spam in it

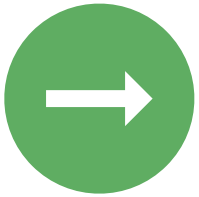
- Character ranges:

RegEx	Text With Match
`[0-9][0-9]`	Player number 42 has taken the field
`[a-zA-Z]ar`	Far, he went, to find a bar that wasn't subpar

MORE LANGUAGE

- Regexes have special meta-characters that represent a type of character
- Decimal digits, whitespace, words, etc.

NEXT UP...



Anchoring expressions to the start or end of a string

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    - **a. Plain matching and Class matching**
    - **b. Meta-characters**
    - **c. Anchors**
    - d. Quantifiers
    - e. Grouping
3. Python Regexes:
    - a. The `re` module
    - b. Accessing Groups
    - c. Named Groups
    - d. Substituting, Splitting, and Escaping
    - e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

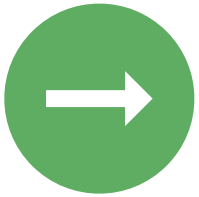
REVIEWING META-CHARACTERS

RegEx	Text With Match
`c.r`	The core of it is, a cur can't drive a car
`b\w\wr`	The bear drank some beer at the bar
`\w\w\w\W\d`	Model numbers were A3F:3, JHK\2, and X23M/1B
`ark\s\w\w`	The ark was large . Inside it was dark and scary.
`\w:\\\w\w\w\\\d\d\d`	Look in folder C:\src\123code\

REGEX LANGUAGE

- You can specify that a regex be part of the beginning or end of a string using an **anchor**
- Multi-line mode changes how some anchors behave around new-lines

NEXT UP...



Repeating expressions with quantifiers

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. **Anchors**
    d. **Quantifiers**
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REVIEWING ANCHORS

RegEx		Text With Match
`^The`		The quick brown fox jumped over The lazy yellow dog
`^The`	Multiline	The quick brown fox jumped over The lazy yellow dog
`\AThe`	Multiline	The quick brown fox jumped over The lazy yellow dog
`red$`		Roses are red Violets are also red
`red$`	Multiline	Roses are red Violets are also red
`red\Z`	Multiline	Roses are red Violets are also red

REVIEWING ANCHORS

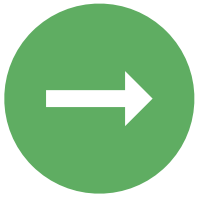
RegEx	Text With Match
`car\b`	The car had nice carpet
`car\B`	The car had nice carpet

REGEX LANGUAGE

- Repetition in patterns is done with quantifiers
- Quantifiers come in four varieties:

`*`	Zero or more
`+`	One or more
`?`	Zero or one
`{m,n}`	Between *m* and *n*, inclusive

NEXT UP...



Grouping parts of expressions

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    - **a. Plain matching and Class matching**
    - **b. Meta-characters**
    - **c. Anchors**
    - **d. Quantifiers**
    - **e. Grouping**
3. Python Regexes:
    - a. The `re` module
    - b. Accessing Groups
    - c. Named Groups
    - d. Substituting, Splitting, and Escaping
    - e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REVIEWING QUANTIFIERS

RegEx	Text With Match
67*3	677734 946773 546734 8463523
67+3	677734 946773 546734 8463523
67?3	677734 946773 546734 8463523

REVIEWING QUANTIFIERS

RegEx		Text With Match
`<.*>`	Greedy	one <two> three <four> five
`<.*?>`	Not Greedy	one <two> three <four> five
`-\d+`	Greedy	93248X-77Y-48348C-5543
`-\d+?`	Not Greedy	93248X-77Y-48348C-5543
`Aa?`	Greedy	Aaaaaah
`Aa??`	Not Greedy	Aaaaaah

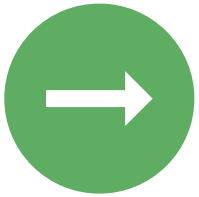
REVIEWING QUANTIFIERS

RegEx		Text With Match
`B{3}`		ABBA ABBB C{A}AB BBBB
`B{2,4}`		ABBA ABBB C{A}AB BBBB
`C{A}A`		ABBA ABBB C{A}AB BBBB
`B{3,4}?`	Not Greedy	ABBA ABBB C{A}AB BBBB

REGEX LANGUAGE

- Parts of a pattern can be grouped together
- Access portions of a match
- Repeat sequences
- Back references

NEXT UP...



Python!

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    - **a. Plain matching and Class matching**
    - **b. Meta-characters**
    - **c. Anchors**
    - **d. Quantifiers**
    - **e. Grouping**
3. **Python Regexes:**
    - **a. The `re` module**
    - b. Accessing Groups
    - c. Named Groups
    - d. Substituting, Splitting, and Escaping
    - e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

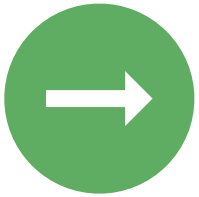
REVIEWING GROUPS

RegEx	Text With Match	
`([aeiou])r`	The earth's core is a long way under us	Match 1: a Match 2: o Match 3: e
`([4-9])+`	45200 2229 8880 2202	Match 1: 5 Match 2: 9 Match 3: 8
`(the).*\1`	the train went down the valley	Match 1: the

REGEXES IN PYTHON

- Regular expressions are in the `re` standard Python library
- Most methods take a pattern string and a string to search against
- Most methods return an `re.Match` object giving information about the match
- `re.Match` objects are “truthy”, they can be compared as booleans

NEXT UP...



Grouped results

TABLE OF CONTENTS

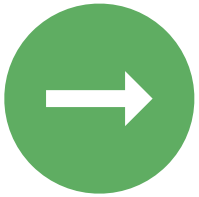
1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. **Anchors**
    d. **Quantifiers**
    e. **Grouping**
3. **Python Regexes:**
    a. **The `re` module**
    b. **Accessing Groups**
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

ACCESSING GROUPS

- The `re.Match` object contains attributes and methods providing details about a match

`.group(*args)`	Returns the matched group or a tuple of matched groups
`.groups()`	Returns all matched groups
`.start()`	Starting index of match
`.end()`	Ending index of match
`.expand(template)`	Returns a string based on template substituting back-references
`.span()`	Returns a tuple containing the start and ending index

NEXT UP...



Named groups

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. **Anchors**
    d. **Quantifiers**
    e. **Grouping**
3. **Python Regexes:**
    a. **The `re` module**
    b. **Accessing Groups**
    c. **Named Groups**
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

BACKSLASHES

- Python strings use backslash for special characters

```
"This has a \t tab and \n newline in it"
```

- To represent a backslash, you have to escape it

```
"This: \\ is only a single backslash"
```

- Regexes also use backslash for meta-characters

```
re.search("\\w+,", "one,two,three")    # escape the backslash to match “one,”
```

- Searching for an actual backslash requires it to be escaped twice

```
re.search("\\\\", "one\\two")    # escaped \\ for Python, then \\ for regex
```

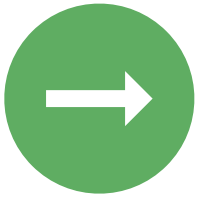
- Raw strings are easier, no escaping

```
re.search(r"\w+\\\w+", r"one\two")    # no escape this time
```

MORE REGEX LANGUAGE

- Python's regex groups also support:
  - Named groups
  - Non-capturing groups

NEXT UP...



More `re` methods

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. **Anchors**
    d. **Quantifiers**
    e. **Grouping**
3. **Python Regexes:**
    a. **The `re` module**
    b. **Accessing Groups**
    c. **Named Groups**
    d. **Substituting, Splitting, and Escaping**
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REVIEWING NAMED & NON-CAPTURING GROUPS

RegEx	Text With Match	
`(?P<digits>\d+)`	I live at 123 Somewhere Street	Match 1: digits: 123
`(?P<digits>\d+).*(?P=digits)`	In some cultures 13 is unlucky, in others 13 is good luck	Match 1: digits: 13
`(a)(?:r)(e)`	I don't care if you don't	Match 1: 1. a 2. e

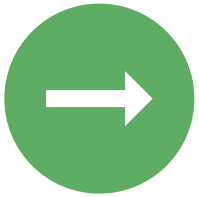
MORE PYTHON REGEX FUNCTIONS

- Python's regex matching functions so far:
  - `re.search()`
  - `re.match()`
  - `re.fullmatch()`
  - `re.findall()`
  - `re.finditer()`

MORE PYTHON REGEX FUNCTIONS

- Substitution:
  - `re.sub()`
  - `re.subn()`
- Utility functions:
  - `re.split()`
  - `re.escape()`

NEXT UP...



Fun with flags

TABLE OF CONTENTS

1. **Overview**
2. **Regular Expressions are a Language:**
    a. **Plain matching and Class matching**
    b. **Meta-characters**
    c. **Anchors**
    d. **Quantifiers**
    e. **Grouping**
3. **Python Regexes:**
    a. **The `re` module**
    b. **Accessing Groups**
    c. **Named Groups**
    d. **Substituting, Splitting, and Escaping**
    e. **Flags**
4. Advanced Matching
5. Fun and Further Reading
6. Summary

FLAGS FOR FINE TUNING

- Most of Python's regex methods support additional flags
- Flags change the behavior of a regex; example: make them case insensitive
- All of these support flags:
  - `re.search()`
  - `re.match()`
  - `re.fullmatch()`
  - `re.findall()`
  - `re.finditer()`
  - `re.sub()`
  - `re.subn()`
  - `re.split()`
  - `re.compile()`

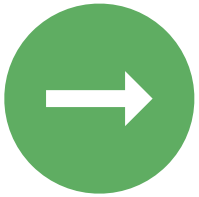
FLAGS

Short Flag	Long Flag	Purpose
`re.I`	`re.IGNORECASE`	Makes matching of alphabetic characters case insensitive
`re.M`	`re.MULTILINE`	Causes ^ and $ anchors to match embedded newlines
`re.S`	`re.DOTALL`	Causes dot meta-character to match a newline
`re.X`	`re.VERBOSE`	Allows whitespace and comments within regex
	`re.DEBUG`	Debug information is printed to console

CHARACTER ENCODING FLAGS

Short Flag	Long Flag	Purpose
`re.A`	`re.ASCII`	Use ASCII encoding for character classification
`re.U`	`re.UNICODE`	Use Unicode encoding for character classification (default in Python 3)
`re.L`	`re.LOCALE`	Use current locale to determine encoding for character classification

NEXT UP...



Trickier matching!

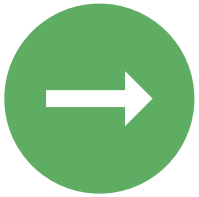
TABLE OF CONTENTS

1. Overview
2. Regular Expressions are a Language:
    a. Plain matching and Class matching
    b. Meta-characters
    c. Anchors
    d. Quantifiers
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

EVEN MORE REGEX LANGUAGE

- Conditional matches
- Look ahead matches
- Look behind matches
- Comments

NEXT UP...



Fun (!?!?)

TABLE OF CONTENTS

1. Overview
2. Regular Expressions are a Language:
    a. Plain matching and Class matching
    b. Meta-characters
    c. Anchors
    d. Quantifiers
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

FUN (!?!?)

- RFC 822 in Perl

http://www.ex-parrot.com/~pdw/Mail-RFC822-Address.html

```
r'(?:(?:\r\n)?[ \t])*(?:(?:(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])?(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])?(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*)*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*\>(?:(?:\r\n)?[ \t])*)(?:,\s*(?:(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*)*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*\<(?:(?:\r\n)?[ \t])*(?:@(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:,@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])?(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)*:(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|"(?:[^\"\r\\]|\\.|(?:(?:\r\n)?[ \t]))*"(?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*)*@(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)(?:\.(?:(?:\r\n)?[ \t])*(?:[^()<>@,;:\\".\[\] \000-\037\177]+(?:(?:(?:\r\n)?[ \t])+|\Z|(?=[\["()<>@,;:\\".\[\]]))|\[(?:[^\[\]\r\\]|\\.)*\](?:(?:\r\n)?[ \t])*)*\>(?:(?:\r\n)?[ \t])*)*)?;\s*)'
```

NOT HAND WRITTEN

```
sub make_rfc822re {
    my $specials = '()<>@,;:\\\\".\\[\\]';
    my $controls = '\\000-\\037\\177';
    my $dtext = "[^\\[\\]\\r\\\\]";
    my $domain_literal = "\\[(?:$dtext|\\\\.)*\\]$lwsp*";
    my $quoted_string = "\"(?:[^\\\"\\r\\\\]|\\\\.|$lwsp)*\"$lwsp*";
    my $atom = "[^$specials $controls]+(?:$lwsp+|\\Z|(?=[\\[\"$specials]))";
    my $word = "(?:$atom|$quoted_string)";
    my $localpart = "$word(?:\\.$lwsp*$word)*";
    my $sub_domain = "(?:$atom|$domain_literal)";
    my $domain = "$sub_domain(?:\\.$lwsp*$sub_domain)*";
    my $addr_spec = "$localpart\@$lwsp*$domain";
    my $phrase = "$word*";
    my $route = "(?:\@$domain(?:,\@$lwsp*$domain)*:$lwsp*)";
    my $route_addr = "\\<$lwsp*$route?$addr_spec\\>$lwsp*";
    my $mailbox = "(?:$addr_spec|$phrase$route_addr)";
    my $group = "$phrase:$lwsp*(?:$mailbox(?:,\\s*$mailbox)*)?;\\s*";
    my $address = "(?:$mailbox|$group)";
    return "$lwsp*$address";
}
```

DIVIDE BY 7

https://github.com/matthiasgoergens/Div7

```
((((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|
1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|
2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|
9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|
9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|4(2|9)*5|7|0)|(((6(2|9)*8|6(2|
9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|
0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|6(2|9)*5|9|2)((((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|
0(2|9)*7|0(2|9)*0|5)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|7(2|9)*6|0(2|9)*6|4)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|
7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|
9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|
6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|4(2|9)*5|7|0)|((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|7(2|9)*7|7(2|9)*0|0(2|9)*7|0(2|9)*0|5)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|
9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|7(2|9)*5|0(2|9)*5|3)*((((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
7(2|9)*7|7(2|9)*0|0(2|9)*7|0(2|9)*0|5)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|
9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|7(2|9)*6|0(2|9)*6|4)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|
1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|
9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|4(2|9)*4|6)|((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|
9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|4(2|9)*4|6)|((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|4(2|9)*4|6)|((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|
9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|4(2|9)*4|6)|((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|
1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|
9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|4(2|9)*4|6)|((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|
9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|8(2|9)*4|1(2|9)*4|3)|(7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*4|7|0)|7(2|9)*4|0(2|9)*4|9|2)|((((6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|
6(2|9)*7|6(2|9)*0|4)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|
1(2|9)*6|5)|(6(2|9)*8|6(2|9)*1|5)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|6(2|9)*6|3)(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)
((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(4(2|9)*8|
4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|4(2|9)*5|7|0)|((3(2|9)*8|3(2|9)*1|9|2)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|3(2|9)*7|3(2|9)*0|8|1)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|
3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|
4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|
4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|8(2|9)*6|1(2|9)*6|5)|(4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*6|9|2)|4(2|9)*6|8|1)*(((4(2|9)*8|4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|4(2|9)*7|4(2|9)*0|9|2)
((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*7|5(2|9)*0|3)|8(2|9)*7|8(2|9)*0|1(2|9)*7|1(2|9)*0|6)*((8(2|9)*8|8(2|9)*1|1(2|9)*8|1(2|9)*1|7|0)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|8(2|9)*5|1(2|9)*5|4)|(4(2|9)*8|
4(2|9)*1|3)(5(2|9)*8|5(2|9)*1|4)*(5(2|9)*5|8|1)|4(2|9)*5|7|0)|((7(2|9)*8|7(2|9)*1|0(2|9)*8|0(2|9)*1|6)...
```

REGEX CODE GOLF

- Original version is 25,604 characters long
- Regex code golf:

  `https://codegolf.stackexchange.com/questions/3503/hard-code-golf-regex-for-divisibility-by-7`
- With constraints and the `.NET` library, 103 bytes

IS IT PRIME?

- Regex for determining if a number is prime

```
def is_prime(n):
    return not re.match(r"^.?$|^(..+?)\1+$", "1"*n)
```

- Converts number to uniary format (1, 11, 111, 1111) == (1, 2, 3, 4)
- Left of the *or*: `^.?$`
  - Matches a single digit 0 or 1 times
- Right of the *or*: `^(..+?)\1+$`
  - Tries to match 2 characters, then 3 characters, etc
  - The back-reference will look for duplicates of such a match, if there are duplicates it matches, which means it isn't prime!
- Originally from a Perl hacker known as Abigail
- Illya Gerasymchuk does a deep dive on how this works:
  https://iluxonchik.github.io/regular-expression-check-if-number-is-prime/

MORE MATH

- Xavier Noria has a whole repo on math by regex:
  `https://github.com/fxn/math-with-regexps`

TOOLS

- Online regex editors:

  `https://pythex.org/`
  `https://regex101.com/`
- Matthew Branett's regex library:
  - nested sets, set operations, infinite look-behind

    `https://bitbucket.org/mrabarnett/mrab-regex/src/hg/`
- Alternative: parse library

  `https://github.com/r1chardj0n3s/parse`

FURTHER READING

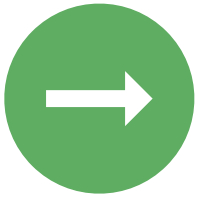
- Wikipedia:

  https://en.wikipedia.org/wiki/Regular_expression
  https://en.wikipedia.org/wiki/Perl_Compatible_Regular_Expressions
  https://en.wikipedia.org/wiki/Finite-state_machine

- Regex Denial of Service (DOS)

  https://levelup.gitconnected.com/the-regular-expression-denial-of-service-redos-cheat-sheet-a78d0ed7d865

NEXT UP...



Summary

TABLE OF CONTENTS

1. Overview
2. Regular Expressions are a Language:
    a. Plain matching and Class matching
    b. Meta-characters
    c. Anchors
    d. Quantifiers
    e. Grouping
3. Python Regexes:
    a. The `re` module
    b. Accessing Groups
    c. Named Groups
    d. Substituting, Splitting, and Escaping
    e. Flags
4. Advanced Matching
5. Fun and Further Reading
6. Summary

REGEX LANGUAGE FEATURES

- Why use a regular expression and where they come from
- Plain matching

```
>>> re.search("spam", "Green eggs and spam")
<re.Match object; span=(15, 19), match='spam'>
```

- Class matching

```
>>> re.search("[0-9]", "I live at the 2nd house")
<re.Match object; span=(14, 15), match='2'>
```

- Meta-characters

```
>>> re.search(“\s\d\w+", "I live at the 2nd house")
<re.Match object; span=(13, 17), match=‘ 2nd'>
```

REGEX LANGUAGE FEATURES

- Anchors

```
>>> re.search("^My", “Hello\nMy name is George", re.MULTILINE)
<re.Match object; span=(6, 8), match='My'>
```

- Quantifiers

```
>>> match = re.search(r"\d+", "My phone number is 555-1212")
<re.Match object; span=(19, 22), match='555'>
```

- Grouping

```
>>> match = re.search(r"(\d+)", "My phone number is 555-1212")
>>> match.groups()
('555',)
```

PYTHON `re` MODULE

- Python's regex matching functions:
  - `re.search()`
  - `re.match()`
  - `re.fullmatch()`
  - `re.findall()`
  - `re.finditer()`

THE `re.Match` OBJECT

- Matching functions return an `re.Match` object
  - `.group()`
  - `.groups()`
  - `.groupdict()`
  - `.expand()`
  - `.start()`
  - `.end()`
  - `.span()`
- Match objects with groups can also be accessed with indexing

```
>>> match = re.search("(\d+)", "My phone number is 555-1212")
>>> match.groups()
('555',)
>>> match[1]
'555'
```

NAMED GROUPS

- Name a group with `(?P<name>regex)`

```
>>> match = re.search(r"(?P<prefix>\d{3})-(?P<lineno>\d{4})", "555-1212")
>>> match.groupdict()
{'prefix': '555', 'lineno': '1212'}
```

- Back reference with `\#` or `(?P=name)`

```
>>> re.search(r"(?P<twice>\d\d)-\1", "314 44-44 94")
<re.Match object; span=(4, 9), match='44-44'>
>>> re.search(r"(?P<twice>\d\d)-(?P=twice)", "314 44-44 94")
<re.Match object; span=(4, 9), match='44-44'>
```

MORE PYTHON REGEX FUNCTIONS

- Substitution:
  - `re.sub()`
  - `re.subn()`
- Utility functions:
  - `re.split()`
  - `re.escape()`

FLAGS

- Flags modify the behavior of the regex

```
>>> re.search("A", "My name is Bob", re.IGNORECASE)
<re.Match object; span=(4, 5), match='a'>
```

- Supported flags:
  - `re.IGNORECASE / re.I`
  - `re.MULTILINE / re.M`
  - `re.DOTALL / re.S`
  - `re.VERBOSE / re.X`
  - `re.DEBUG`
  - `re.ASCII / re.A`
  - `re.UNICODE / re.U`
  - `re.LOCALE / re.L`

ADVANCED MATCHING

- Conditional matches:
  - Changes the matching criteria based on the presence or absence of a group
- Look ahead / behind
  - Grouped match after or before the match
  - Is not included as part of the match

Thanks!

Dankie ju faleminderit faleminderit شكرا Grazias Շնորհակալություն Sağ ol eskerrik asko Дзякуй তোমাকে ধন্যবাদ hvala trugéré
благодаря Akeva Chezu ba gràcies Salamat zikomo 谢谢 hvala děkuji Tak dank u Dankon aitäh takk fyri salamat kiitos Merci
Grazas დიდი მადლობა Danke σας ευχαριστώ આભાર Mèsi poutèt ou Na gode Mahalo תודה Dhanyawaad köszönöm þakka þér
Daalụ terima kasih Go raibh maith agat ありがとう សូមអរគុណអ្នក Kamsahamnida ຂໍຂອບໃຈທ່ານ
Lorem ipsum dolor paldies ačiū Mauruuru Dhanyawaadh Welálin баярлалаа
barka Ahéhee' Dhanyabaad miigwetch manana شکر از شما dziękuję obrigado ਤੁਹਾਡਾ ਧੰਨਵਾਦ mulţumesc спасибо tapadh leibh хвала
ďakujem hvala Waad ku mahadsan tahay Gracias Asante Tack Salamat rahmat நன்றி ధన్యవాదాలు ขอบคุณ tualumba teşekkür
ederim Спасибі آپ کا شکریہ rahmat cảm ơn bạn Diolch yn fawr אַ דאַנק Balika o ṣeun Ngiyabonga