

SPICE WORLD

2022
HYBRID



Unlocking the Power of Regular Expressions with PowerShell

<https://github.com/jdhitsolutions/SpiceWorld2022-PSRegex-Intro>



SPICEWORLD2022
HYBRID



Jeff Hicks

JDH Information Technology Solutions, Inc.
@jeffhicks



SPICEWORLD2022
HYBRID





What is a Regular Expression?

- A way of describing data (text)
- Based on a set of accepted and standardized patterns
- Introduced in 1951
 - Incorporated into Unix
 - Led to *grep*
 - “*Global search for Regular Expressions and Print lines*”
- PowerShell supports the full regular expression library from .NET



Why Regular Expressions

- Validating data
- Finding data
- Fixing data
 - Splitting
 - Replacing

Patterns

- (202) 555-9876
- 192.168.2.3
- \\server01\public





Patterns

- `\(\d{3}\)\s\d{3}-\d{4}`
- `(\d{1,3}\.){3}\d{1,3}`
- `\\\\\\w+\\\\\\w+`



Regular Expressions are Classy

| Pattern | Match |
|---------|--|
| \w | Any word character, including numbers and the underscore (_) |
| \d | Any digit |
| \D | Any non-digit |
| \s | Any whitespace |
| \W | Any non-word character |
| \S | Any non-whitespace |

Case-Sensitive

With Qualifiers

| Qualifier | Description |
|-----------|--|
| * | Match 0 or more of the preceding pattern |
| + | Match 1 or more of the preceding pattern |
| ? | Match 0 or 1 instance of the preceding pattern |
| {n} | Match <i>exactly</i> N number of the preceding pattern |
| {n,m} | Match <i>at least</i> N number of the preceding pattern and <i>no more</i> than M number |
| {n,} | Match <i>at least</i> N number of the preceding pattern |

Special Characters

| Character | Description |
|-----------|--|
| . | Any single character |
| [xyz] | Match at least one of the characters in the brackets |
| [a-zA-Z] | Match at least one of the characters in the range. Case sensitive. |
| [^xyz] | Match any character <i>except</i> what is in brackets |
| ^ | Match the beginning characters |
| \$ | Match the ending characters |
| \ | The regular expression escape character |



Escape Clause

- You'll need to escape regular expression characters if you need to match them
 - `[]<>?^$`
- `\`
- `[2022-12-22]` Some data
 - `\[[\d-]+\]`
 - Match on the literal `[`
 - `\[`
 - Match on a digit or the dash
 - `[\d-]`
 - Match on the literal `]`
 - `\]`

Show Me

<https://github.com/jdhitsolutions/SpiceWorld2022-PSRegex-Intro>



SPICEWORLD2022
HYBRID



Regular Expression “Gotchas”

- Comparisons are True/False
- Watch the float
 - How exact do you need to be?
- Test for known failures
- You still might need to validate
- You have to know your data
 - It should be predictable
 - It should be consistent



Advanced Stuff

- Look ahead and look behind
 - `(?<=\d)\s\w+\.\w+`
- Optional matches
 - `\d{3}(\-[a-z]+)?`
- Regex Options
 - Configuring case
 - Multiline search
- More with the [regex] type accelerator
 - `[regex]$rx = "[a-z]{1}\.\w{2,7}@company.com"`

Resources

[Help about_regular_expressions](#)

[Help about_comparison_operators](#)

[Help about_wildcards](#)



Resources

<https://www.regular-expressions.info>

<https://rubular.com>

<https://regex101.com>

<https://pluralsight.pxf.io/psregex>



Find Me Online

@jeffhicks

<https://jdhitsolutions.com/blog>

<https://github.com/jdhitsolutions>

<https://jeffhicks.substack.com>

<https://pluralsight.pxf.io/jeffhicks>



Thank you.



SPICEWORLD2022
HYBRID