Session 11: Markdown

P. S. Langeslag

10 January 2019

Standard Markdown (`Markdown.pl`)

▶ Released in 2004 as a language *and* a Perl conversion script

▶ Written by John Gruber with help from Aaron Swartz

▶ Provides readable structured plaintext

▶ The conversion script outputs to XHTML (or HTML)

See https://daringfireball.net/projects/markdown/.

Standard Markdown Philosophy

- ▶ A publishing format for the web
- ▶ Source sufficiently plain that it can be published as-is
- ▶ Intuitive formatting using only "punctuation characters"
- ▶ Inspired primarily by plaintext email formatting conventions
- ▶ An informal specification: lenient; permits different styles
- ▶ For complex formatting (tables), use HTML *inside* your Markdown

Shortcomings of Standard Markdown

Syntax

▶ No native table support (HTML source tables are hard to read)
▶ No support for footnotes, smart quotes, small capitals, strikeout, image captions, automated citations, math, LaTeX code

Output Formats

▶ No LaTeX, PDF, Word, OpenText, presentation slides, etc.

Markdown Forks

- ► CommonMark
  - ► A strongly defined, unambiguous Markdown specification

Markdown Forks

- ► **CommonMark**
  - ► A strongly defined, unambiguous Markdown specification

- ► **MultiMarkdown**
  - ► Adds footnotes, smart quotes, tables with colspan, table/image captions, LaTeX, citations, math, additional output formats

Markdown Forks

- ▶ **CommonMark**
  - ▶ A strongly defined, unambiguous Markdown specification

- ▶ **MultiMarkdown**
  - ▶ Adds footnotes, smart quotes, tables with colspan, table/image captions, LaTeX, citations, math, additional output formats

- ▶ **Github-Flavored Markdown (GFM)**
  - ▶ CommonMark with extensions for tables, strikeout, autolists, and task lists

Markdown Forks

- **CommonMark**
  - A strongly defined, unambiguous Markdown specification

- **MultiMarkdown**
  - Adds footnotes, smart quotes, tables with colspan, table/image captions, LaTeX, citations, math, additional output formats

- **Github-Flavored Markdown (GFM)**
  - CommonMark with extensions for tables, strikeout, autolists, and task lists

- **Pandoc Markdown**
  - Adds everything mentioned thus far except colspan
  - Well integrated with LaTeX
  - Pandoc tool converts to a plethora of formats

Markdown Forks

- ▶ CommonMark
  - ▶ A strongly defined, unambiguous Markdown specification

- ▶ MultiMarkdown
  - ▶ Adds footnotes, smart quotes, tables with colspan, table/image captions, LaTeX, citations, math, additional output formats

- ▶ Github-Flavored Markdown (GFM)
  - ▶ CommonMark with extensions for tables, strikeout, autolists, and task lists

- ▶ Pandoc Markdown
  - ▶ Adds everything mentioned thus far except colspan
  - ▶ Well integrated with LaTeX
  - ▶ Pandoc tool converts to a plethora of formats

- ▶ Markdown Extra
  - ▶ Based on a PHP(/Python/Ruby) implementation
  - ▶ Adds footnotes, tables, abbreviations, Markdown inside HTML code, id/class attributes

Seamless Integration

▶ Standard Markdown allows HTML syntax

▶ Pandoc Markdown allows HTML and LaTeX

You can thus use any LaTeX tricks and tweaks, such as `\today`, `\vspace*{3mm}`, `\LaTeX`, etc. Pandoc will preserve what LaTeX functionality it can in non-LaTeX output formats.

(Standard/Pandoc) Markdown Escape Characters

Markdown/HTML/LaTeX special characters are interpreted according to context, but may be backslash-escaped where required:

\	backslash
`	backtick
*	asterisk
_	underscore
{}	curly braces
[]	square brackets
()	parentheses
#	hash mark
+	plus sign
-	minus sign (hyphen)
.	dot
!	exclamation mark

e.g. 3 < 4 and <4> are printed as is, but <alpha> is parsed as HTML.

Pandoc Syntax: Document Structure

```
# Top-Level Heading
```

```
This is a paragraph.
```

```
Paragraphs and other block-level elements (headings, lists,
tables, etc.) are delimited by blank lines.
```

```
## Second-Level Heading
```

```
You can use up to six levels of headings (though you may have
to provide definitions for the lowest levels in your output
format).
```

Line Breaks and Line Blocks

If you need a custom line break, end the line with two or more spaces.

To insert line breaks structurally (e.g. in address blocks or verse), consider the line block, which respects line breaks in the source:

```
Regular paragraph here.

| I have eaten
| the plums
| that were in
| the icebox
|
| and which
| you were probably
| saving
| for breakfast

Next paragraph here.
```

Compiling Pandoc

```
pandoc -o output.pdf input.md --pdf-engine xelatex

pandoc -o output.docx input.md --pdf-engine xelatex

pandoc -o output.pdf input.md --pdf-engine xelatex
 --filter pandoc-citeproc
```

Pandoc Syntax: Bullet Lists

Use -, *, or + as item markers.

```
# Shopping List

Don't forget to grab these essential groceries:

- Tomatoes
- Cucumber
- Lettuce

Else there will be no lunch.
```

Pandoc Syntax: Ordered Lists

```
1. These items
2. will appear
6. numbered consecutively
#. 1, 2, 3, 4.
    #. Lists can be nested by tab-indenting the lower level.

a. This creates a
b. Different marker format

A) And
B) This

(IV) And
(V) This

A.  Capital markers followed by periods
B.  Must be followed by two spaces
```

Pandoc Syntax: Advanced Lists

► For extra vertical space, leave a blank line after each item.

► To restart numbering (multiple subsequent lists with the same marker type), separate them with a comment tag `<!-- -->`.

► Different marker types automatically restart numbering.

► For additional paragraphs, or block quotes within list items, leave a blank line before and after such subelements but line them up with the start of the item content (after the space).

► For code blocks within items, indent two tabs (eight spaces)

Pandoc Syntax: Title Block

```
% Title
% Author One; Author Two; Author Three
% \today
```

- ▶ Place at the very top of your document
- ▶ Later fields (date; or author and date) may be omitted
- ▶ Leave fields blank e.g. to omit the title but include the author
- ▶ Use a leading space for any subsequent lines of output
- ▶ Authors are separated with either a semicolon or a line break and leading space

Pandoc Syntax: Character Formatting

▶ _italics_ *or* *italics*
▶ __bold__ *or* **bold**
▶ ~~strikeout~~
▶ ^superscript^
▶ ~subscript~
▶ [text here]{.smallcaps}
▶ `verbatim/monospace`
▶ ``verbatim/monospace containing a backtick``
▶ `verbatim with syntax highlighing`{.xml}
  (see pandoc --list-highlight-languages for options)

Pandoc Syntax: Block Quotations

```
Block quotations follow a greater-than character:
```

```
> Lorem ipsum dolor sit amet, consectetur adipiscing
> elit. Cras sed enim molestie, fringilla eros a,
> sodales elit. Fusce eu quam sit amet eros ornare
> ultricies.
```

A wrapped line only requires a single >, but you may choose to hard-break your lines for aesthetic reasons and add one > per line.

Pandoc Syntax: Block Quotations

Block quotations follow a greater-than character:

> Lorem ipsum dolor sit amet, consectetur adipiscing
> elit. Cras sed enim molestie, fringilla eros a,
> sodales elit. Fusce eu quam sit amet eros ornare
> ultricies.

A wrapped line only requires a single >, but you may choose to hard-break your lines for aesthetic reasons and add one > per line.

Nest as follows:

> This block quote contains a further block quote:
>
> > This is the contained block quote.

Pandoc Syntax: Verbatim/Code Blocks

Indented Code Blocks

Set off by four spaces or one tab space:

```
if [ -z $DISPLAY ]; then
    printf "Press D to start DWM or any other key to start
     XMonad"
fi
```

Normal text continues here.

Pandoc Syntax: Verbatim/Code Blocks

Fenced Code Blocks

Set off by three backticks/tildes and an optional language ID.
Short form:

```bash
if [ -z $DISPLAY ]; then
    startx
fi
```

In long format, fenced code blocks take further attributes:

Set off by three backticks or tildes and any desired attributes:

``` {.bash .numberLines}
grep -io 'pattern' filename
```

Pandoc Syntax: Comments

Use HTML comments:

```
<!-- This text is ignored by the parser -->
```

Pandox Syntax: Hyperlinks

```
[Link Text](https://url.org) <!-- An "inline link" -->
[Link Text](https://url.org "Optional Title")
[Link Text](mailto:user@domain.org)
<https://url.org> <!-- An "automatic link" -->
<user@domain.org>
```

▶ Automatic links autodetect email addresses, inline links do not
▶ Titles are optional
▶ Link text allows formatting; titles do not

Pandoc Syntax: Reference Links

Place a (list of) link definition(s) anywhere in your document, then refer to them by a label.

Explicit Reference Links

```
[label]: https://url.org "Optional title"
```

```
Here I cite [the resource defined above][label].
```

Implicit Reference Links

```
[label]: https://url.org "Optional title"
```

```
Here I cite [label][].
```

Shortcut Reference Links

```
[label]: https://url.org "Optional title"
```

```
Here I cite [label].
```

Pandoc Syntax: Internal Links

Link IDs are generated from headings:

```
# My Heading

Some text

# My Heading
```

As was seen in [the previous section](#my-heading-1), . . .

See the Pandoc User's Guide for details on the algorithm.

Pandoc Syntax: Images

```
![Caption goes here](img/filename.png)
```

Explicit scaling is done as follows, referencing `width` or `height`:

```
![Caption goes here](img/filename.png){ width="25%" }
```

Pandoc Syntax: Basic Tables
Headings and Full Caption Syntax

```
Column 1 Heading    Column 2 Heading   Column 3 Heading
------------------  -----------------  ------------------
Row 1 column 1      Row 1 column 2     Row 1 column 3
Row 2 column 1      Row 2 column 2     Row 2 column 3

Table: Caption goes here
```

Pandoc Syntax: Basic Tables

Headings and Full Caption Syntax

```
Column 1 Heading    Column 2 Heading    Column 3 Heading
------------------  ------------------  ------------------
Row 1 column 1      Row 1 column 2      Row 1 column 3
Row 2 column 1      Row 2 column 2      Row 2 column 3


Table: Caption goes here
```

No Headings and Shorthand Caption

```
--------------  --------------  ---------------
Row 1 col. 1    Row 1 col. 2    Row 1 col. 3
Row 2 col. 1    Row 2 col. 2    Row 2 col. 3
--------------  --------------  ---------------

: Caption goes here
```

▶ Rightmost column content may cross the source table boundary.
▶ Bottom dashed lines need not reflect column width.

Pandoc Syntax: Column Alignment

Heading alignment relative to the line column determines alignment
for the whole column:

```
Left              Centre              Right
-------------    --------------    ---------------
left-aligned     centred           right-aligned

Table: Caption goes here
```

Pandoc Syntax: Multiline Tables

```
-------------------------------------------------
Column 1                    Column 2
------------------------    ---------------------
In this syntax, cell        I find myself using
content may take up         this format mostly.
multiple lines.

Leave a blank line to       All Markdown tables
start a new row.            are a pain to modify.
-------------------------------------------------

Table: Use this syntax to allow multiline cell content
```

Rightmost column content may cross the source table boundary and even form a wrapped line.

Pandoc Syntax: Multiline Tables, No Heading or Caption

```
-------------------------  ---------------------
Multiline tables can       Captions are optional,
likewise do without        too.
headings.

Remember to leave a blank   The bottom dashed line
line between any two rows.  can be continuous.
--------------------------------------------------

: This is the shorthand caption syntax again.
```

Pandoc Syntax: Troubleshooting Tables

▶ Make sure your line columns are wider than your content (not required of rightmost column)

▶ Tweak relative column width using line column width

▶ Check that your syntax matches the chosen table type:

Pandoc Syntax: Troubleshooting Tables

▶ Make sure your line columns are wider than your content (not required of rightmost column)

▶ Tweak relative column width using line column width

▶ Check that your syntax matches the chosen table type:

  ▶ Basic table with header row
    ▶ Minimum of one dashed-line rule indicating columns
    ▶ No blank line between rows
    ▶ End with blank line or dashed line followed by blank line

Pandoc Syntax: Troubleshooting Tables

▶ Make sure your line columns are wider than your content (not required of rightmost column)

▶ Tweak relative column width using line column width

▶ Check that your syntax matches the chosen table type:

   ▶ Basic table with header row
     ▶ Minimum of one dashed-line rule indicating columns
     ▶ No blank line between rows
     ▶ End with blank line or dashed line followed by blank line
   ▶ Basic table without header row
     ▶ Dashed-line rule top and bottom, top indicates columns

Pandoc Syntax: Troubleshooting Tables

- ▶ Make sure your line columns are wider than your content (not required of rightmost column)

- ▶ Tweak relative column width using line column width

- ▶ Check that your syntax matches the chosen table type:
  - ▶ Basic table with header row
    - ▶ Minimum of one dashed-line rule indicating columns
    - ▶ No blank line between rows
    - ▶ End with blank line or dashed line followed by blank line
  - ▶ Basic table without header row
    - ▶ Dashed-line rule top and bottom, top indicates columns
  - ▶ Multiline table with header row
    - ▶ Three dashed-line rules, the second of which indicates colums
    - ▶ Blank line between rows
    - ▶ End with blank line after bottom dashed line

Pandoc Syntax: Troubleshooting Tables

- ▶ Make sure your line columns are wider than your content (not required of rightmost column)

- ▶ Tweak relative column width using line column width

- ▶ Check that your syntax matches the chosen table type:
  - ▶ Basic table with header row
    - ▶ Minimum of one dashed-line rule indicating columns
    - ▶ No blank line between rows
    - ▶ End with blank line or dashed line followed by blank line
  - ▶ Basic table without header row
    - ▶ Dashed-line rule top and bottom, top indicates columns
  - ▶ Multiline table with header row
    - ▶ Three dashed-line rules, the second of which indicates colums
    - ▶ Blank line between rows
    - ▶ End with blank line after bottom dashed line
  - ▶ Multiline table without header row
    - ▶ Two dashed-line rules, the first of which indicates columns

Pandoc Syntax: Footnotes

```
Arguably the  most user-friendly footnote feature
is `inline_notes`.^[This is the footnote.]
```

Columns

Cf. the LaTeX `multicols` and `minipage` environments

```
::::::::::::: {.columns}
::: {.column width="50%"}
Left-hand column content goes here.
:::
::: {.column width="50%"}
Right-hand column content goes here.
:::
:::::::::::::
```

YAML Metadata Blocks

- ▶ Allows global configuration otherwise entered as command-line options
- ▶ Set off by three hyphens top and bottom (leave a blank line above the block if it does not occur at the very top of the document)

```
---
fonttheme: serif
mainfont: Junicode
monofont: "DejaVu Sans Mono"
highlight-style: tango
colorlinks: true
csl: chicago-fullnote-bibliography-with-ibid.csl
bibliography: slides.bib
reference-section-title: "Bibliography"
nocite: |
  @pandoc, @markdown, @multimarkdown, @commonmark, @markdownguide
---
```

The YAML Format

- ▶ `key: value` (the space is mandatory)
- ▶ If the value contains spaces or colons, quote the full value:
  `mainfont: "Andron Scriptor Web"`
- ▶ Arrays take one key and value pair per line, starting with a hyphen:

```
author:
- Jane F. Doe
- John F. Doe
```

- ▶ Fields with blank lines or block-level formatting must be so identified using a pipe:

```
abstract: |
    Paragraph 1.

    Paragraph 2.
```

Selected Global YAML Metadata Variables

Not all are available for all output formats.

▶ `author`, `title`, `date` (if not using title block)

▶ `subtitle` (for output formats/classes that accept it)

▶ `abstract`

▶ `toc: true` (generate a table of contents)

▶ `toc-depth: 2` (specify what levels are represented in the TOC)

▶ `bibliography`

▶ `reference-section-title` (bibliography heading)

▶ `csl` (identifies a stylesheet specification)

▶ `figPrefix:`
      `- "Figure"`
      `- "Figures"`

▶ `tblPrefix:`
      `- "Table"`
      `- "Tables"`

Selected YAML Metadata Variables for LaTeX

- `papersize`: `a4` (or `letter`, etc.)

- `fontsize`: `12pt` (or `10pt`)

- `fonttheme`: `serif` (for `beamer` slides)

- `mainfont`, `sansfont`, `monofont`

- `linestretch`: `1.5` (line height)

- `indent`: `true` (use first-line indent)

- `parskip`: `0pt` (no vertical space between paragraphs)

- `colorlinks=true` (for `hyperref`)

- `header-includes` for header contents (e.g. to load packages)

Citation Commands in Brief

```
Citations are produced thus.[@source1; @source2 55-60;
 @source3 chapter 2]

This suppresses the author's name.[-@johndoe 10-15]

According to @source1 [20], this is an inline citation.
```

Compiling with `biblatex`

Compile from the command line using `--filter pandoc-citeproc`.

Recommended Reading

"CommonMark." Accessed January 6, 2019.
https://commonmark.org.

Cone, Matt. *Markdown Guide*. Accessed January 6, 2019.
https://www.markdownguide.org.

Gruber, John. "Markdown." Accessed January 6, 2019.
https://daringfireball.net/projects/markdown/.

MacFarlane, John. "Pandoc User's Guide." Accessed January 6, 2019.
http://pandoc.org/MANUAL.html.

Penney, Fletcher T. "MultiMarkdown." Accessed January 6, 2019.
https://fletcherpenney.net/multimarkdown/.