Using pandoc for scientific markdown

This is a little tutorial I wrote for myself to help me get started with using markdown for scientific documents and as a resource for when I forget how to do stuff. I've stuck it here in case it turns out to be useful for anyone else, but I'm not promising anything.

As ever, this is a collection of things I found out by frantic googling, so this tutorial is indebted to all the tutorials and stack overflow questions already out there.

Code

Code can be specified inline using backticks like this: `` `some code` ``, which renders this: `some code`. Code blocks can be rendered by tab indenting a paragraph, so this:

```
    a
    code
    block
```

is rendered as this:

```
a
code
block
```

To do an *indented* code block use two tabs, and so on.

Equations

Equations can be specified using LaTeX code like this:

```
    $\pi^2$
```

which renders as π^2.

Equations can be numbered too, by adding a reference in the text and an anchor to the maths like this:

```
equation (@first):

(@first) $\pi^2 + 2 - 3^n$
```

which renders as:

equation (1):

(1) $\pi^2 + 2 - 3^n$

Figures

Figures can be included similarly to urls using the format:

```
![Figure caption][/path/to/figure.fig]
```

or alternatively with a reference and the path at the end, like this:

```
![Figure caption][id]
...
[id]: path/to/figure.fig
```

For example, the following line of code

```
![**Fig. 1**: a bunch of points or something][points]
```

(with the file path at the end of the doc, referenced by `[points]`) returns this figure and caption:

Unfortunately, the filepaths go a bit funny if outputting to docx and the filepath is anywhere other than in the same directory as the output file or below.

Also, LaTeX adds the 'Figure n' bits in itself, so the above will have that repeated in the pdf.

Figure referencing

Unfortunately, explicitly referencing figures by number is harder. If we're rendering to pdf we can use LaTeX references of the form `\ref{id}` in the text and `\label{id}` in the figure caption. This doesn't work for formats like html and docx though.

So markdown like this:

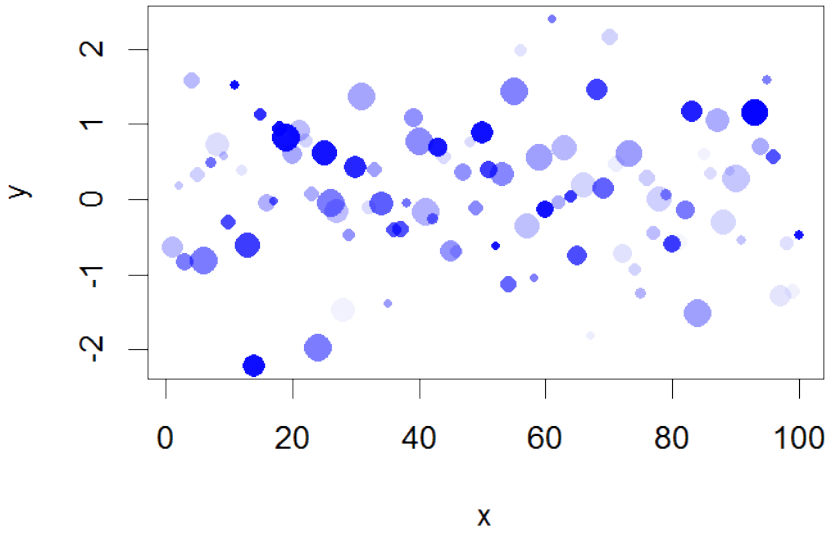

Figure 1: **Fig. 1**: a bunch of points or something

```
Figure \ref{bar} is great.

![**Figure 2**: some pretty bars \label{bar}][bars]
```

(again with the filepath at the end) renders correctly in pdf, but not so well in html or docx:

Figure 2 is great.

Citations

Citations are specified by citekeys in the format `@Ward2009`, which is rendered as: Ward et al. (2009).

To fill in the citations and add a bibliography (at the end of the document, `pandoc` needs us to specify a `bibtex` bibliography file. We can also change the citation style by specifying a `.csl` file. These are both covered in the 'pandoc' section below.

pandoc

Assuming `pandoc` is installed and in the system path, rendering the file is as simple as running the following from the command line:

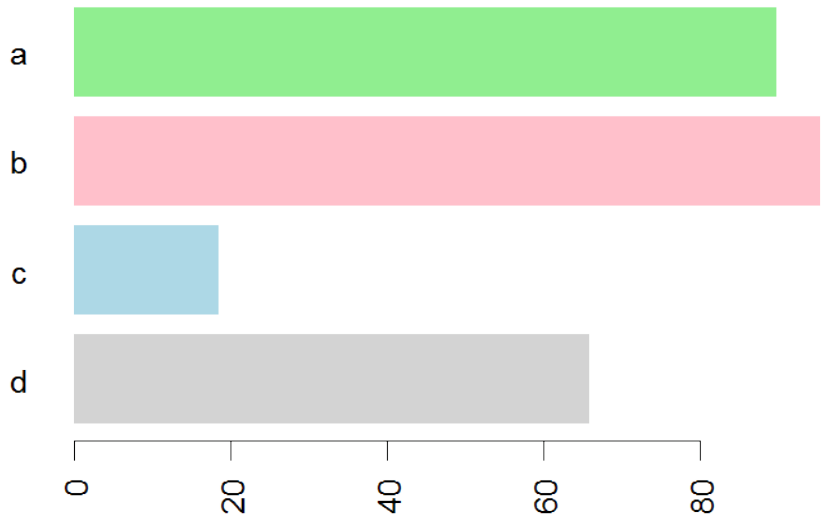

Figure 2: **Figure 2**: some pretty bars

```
pandoc file.md -o file.html
```

To get all the useful scientific features working, we need a bunch of other options:

- building a standalone file (`-s`)
- using smart quotes (`-S`)
- rendering the file using pandoc-flavoured markdown (`-c pandoc.css`)
- specifying a `.bib` file to render citations (`--bibliography=file.bibtex`)
- specifying a `.csl` file to set the citation style (`--csl=file.csl`)

giving us something like this:

```
    pandoc -s -S -c pandoc.css --bibliography=file.bibtex --csl=file.csl file.md -o file.htm
```

or

```
    pandoc -s -S -c pandoc.css --bibliography=file.bibtex
        --csl=file.csl file.md -o file.docx
```

So to render this file I used

```
pandoc -s -S -s pandoc.css --bibliography=library.bib
    --csl=mee.csl test.md -o docs/test.html

pandoc -s -S -s pandoc.css --bibliography=library.bib
    --csl=mee.csl test.md -o docs/test.docx

pandoc -s -S -s pandoc.css --bibliography=library.bib
    --csl=mee.csl test.md -o docs/test.pdf
```

References

Ward, G., Hastie, T., Barry, S., Elith, J. & Leathwick, J.R. (2009). Presence-only data and the em algorithm. *Biometrics*, **65**, 554–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18759851