STATS 531 project report format

Abstract

This is the abstract. It should briefly describe the context of the report and highlight the findings of highest interest to the target readers (other students in the class, the GSI, and the instructor). Usually, the abstract is 150-300 words.

1 Introduction

The source code for this document provides a template based on Quarto for STATS 531, We demonstrate `Python` (Perez, Granger, and Hunter 2011) code with some discussion of how `R` (R Core Team 2020) code can be included.

2 Methods

This section covers basic markdown techniques.

2.1 Formatting

Quarto is a versatile formatting system for authoring HTML based on markdown, integrating LaTeX and various code block interpreted either via Jupyter or Knitr (and thus deal with Python, R and many other langages). It relies on the Pandoc Markdown markup language.

To render/compile a document, run `quarto render`. A document will be generated that includes both content as well as the output of any embedded code chunks within the document. To generate `template.pdf`,

```
quarto render template.qmd
```

2.2 Basic markdown

Bold text or *italic*

- This is a list
- With more elements
- It isn't numbered.

But we can also do a numbered list

1. This is my first item
2. This is my second item
3. This is my third item

In scientific reports, bullet points are seldom used.

2.3 Mathematics

LaTeX code is natively supported, which makes it possible to use mathematical formulae:

$$f(x_1, \dots, x_n; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right) \tag{1}$$

It is also posible to cross-reference an equation, see Equation 1 and Equation 2 for examples.

$$\begin{aligned} D_{x_N} &= \frac{1}{2}\begin{bmatrix} x_L^\top & x_N^\top \end{bmatrix}\begin{bmatrix} L_L & B \\ B^\top & L_N \end{bmatrix}\begin{bmatrix} x_L \\ x_N \end{bmatrix} \\ &= \frac{1}{2}(x_L^\top L_L x_L + 2x_N^\top B^\top x_L + x_N^\top L_N x_N), \end{aligned} \tag{2}$$

Numbering and cross-referencing makes your report easier to read and easier to peer review. That is a consideration for good scholarship

3 Results

A reproducible document should not contain hard-coded results. Instead, it should run code to generate these results from data. Quarto uses either Jupyter or knitr to render code chunks. This can be triggered in the yaml header, e.g., for Jupyter (should be installed on your computer) use

```
---
title: "My Document"
author "Jane Doe"
jupyter: python3
---
```

Here, we use Jupyter with Python. As an example of a simple `Python` chunk:

```
import numpy as np
np.random.seed(0)
x = np.random.normal(size=10)
```

Variables can be accessed inline, outside of code chunks, for example, the mean of the variable `x` is 0.738.

To incorporate some R, called from Python, you can use `%%R` cell magic via the `rpy2` package. Alternatively, you can run the qmd file via R and include some Python via the `reticulate` package (Ushey, Allaire, and Tang 2020). It may be simplest to use a single language, and here we stick to Python.

3.1 Figures

```
import matplotlib.pyplot as plt
plt.figure(figsize=(8,3))
plt.plot(range(10),x)
plt.tight_layout(); plt.show()
```

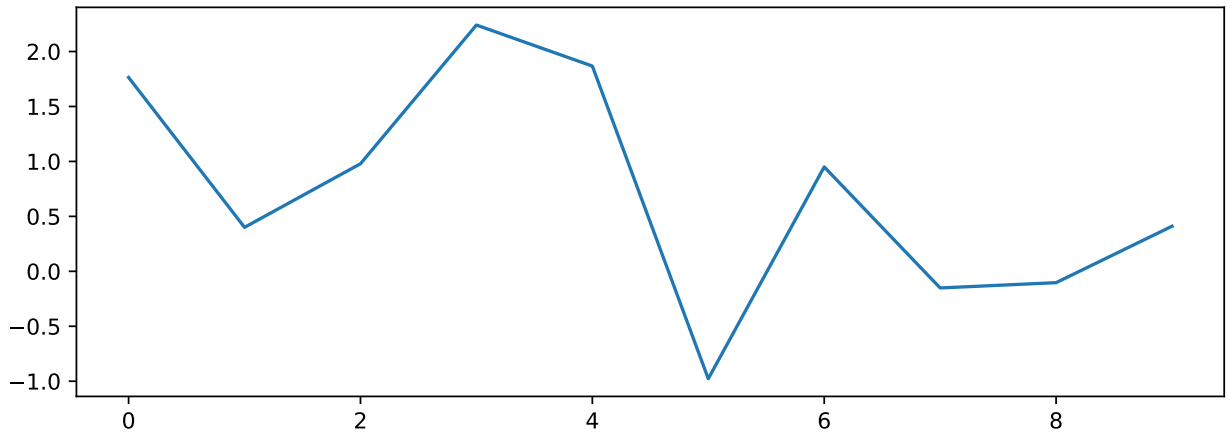

Figure 1: A time plot of x

Plots can be generated and referenced as shown. See plot Figure 1. Usually, we set `echo: false` in the code chunk to suppress the code in the report. The qmd source should be available for those who want to check the code.

3.2 Tables

Tables can be generated with markdown as follows. The label, `@tbl-mylabel`, renders as Table 1.

```
| Tables   |      Are      |  Cool |
|----------|:-------------:|------:|
| col 1 is |  left-aligned | $1600 |
| col 2 is |    centered   |   $12 |
| col 3 is | right-aligned |    $1 |
: my table caption {#tbl-mylabel}
```

Table 1: my table caption

Tables	Are	Cool
col 1 is	left-aligned	$1600
col 2 is	centered	$12
col 3 is	right-aligned	$1

For tabulating results of a data analysis, it is useful to know that Quarto automatically renders the DataFrame as a markdown table when it is the last line of the chunk:

```
import pandas as pd
data = {
    'Fruit': ['Apple', 'Pear', 'Orange'],
    'Price': [2.05, 1.37, 3.09],
    'Quantity': [100, 150, 80]
}

df = pd.DataFrame(data)
df
```

Table 2: A summary table of example data.

	Fruit	Price	Quantity
0	Apple	2.05	100
1	Pear	1.37	150
2	Orange	3.09	80

4 Conclusions

As well as discussing the main results, this section can include speculative interpretation. References are welcome to connect your analysis to the broader literature.

4.1 Bibliographic references

References are displayed using BibTeX, e.g. `[@computo]` will be displayed as (Computo Team 2021), where `computo` is the bibtex key for this specific entry. `@computo` gives an unparenthesized reference to Computo Team (2021). The bibliographic information is automatically retrieved from the `.bib` file specified in the header of this document (here:`references.bib`). Some BibTeX entries relevant to STATS 531 are already in `references.bib`, including Huang and Petukhina (2022), Ionides (2026), Shumway and Stoffer (2017).

4.2 Other cross-references

As already (partially) seen, Quarto includes a mechanism similar to the bibliographic references for sections, equations, theorems, figures, lists, etc. Have a look at this page. For example, you may want to cross-reference to the supplementary material (Section 5).

Acknowledgments

This template builds on a template document by Shao-Ting Chiu. AI was used for debugging via Google AI overview. Vscode with the Quarto extension was used for editing.

Bibliography

Computo Team. 2021. "Computo: Reproducible Computational/Algorithmic Contributions in Statistics and Machine Learning." *Computo*.

Huang, Changquan, and Alla Petukhina. 2022. *Applied Time Series Analysis and Forecasting with Python*. Springer.

Ionides, Edward. 2026. "Notes for STATS 531, Modeling and Analysis of Time Series Data." https://ionides.github.io/531w26/.

Perez, Fernando, Brian E Granger, and John D Hunter. 2011. "Python: An Ecosystem for Scientific Computing." *Computing in Science & Engineering* 13 (2): 13–21.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Shumway, Robert H, and David S Stoffer. 2017. *Time Series Analysis and Its Applications: With R Examples*. 4th ed. Springer.

Ushey, Kevin, JJ Allaire, and Yuan Tang. 2020. *Reticulate: Interface to Python*. https://github.com/rstudio/reticulate.

5 Supplementary material

You can put additional analysis here, or other relevant material. You should not assume the reader will read the supplement systematically. Material in the supplement should support claims made in the body of the report, and should be referenced in the report. The strongest and most critical evidence should be in the body of the report, not in the supplement.