Day 1 - R and regression review

Introduction

> ChatGPT: What is time series analysis?
>
> The study of time series is the study of statistical techniques used to analyzed time-ordered data points. The goal is to identify patterns, trends, and relationships within the data over time. This type of analysis is crucial in various fields, including finance, economics, environmental science, and engineering. Key components and objectives of time series analysis include: trend analysis, understanding seasonality, forecasting, and decomposition.

The study of serially correlated data builds upon the basic framework of linear models. The purpose of today's lab/lecture is to ensure everyone is both proficient with the software required for this course and with linear models and regression.

A quick note on separating your response from the rest of the text in assignments: when possible, please use callout boxes to separate your responses from the rest of the text, such as:

> Here is the answer!

R, RStudio, and TinyTex

1. Install R and RStudio. R and RStudio may be installed for free at https://cran.r-project.org/ and https://posit.co/downloads/, respectively.

   > Note
   >
   > RStudio is an integrated development environment (IDE) that we will use to interact with the statistical software R. Seldom will we interact with the program R directly.

2. RStudio (but really, Pandoc) allows us to render a combination of plain text, code, and output into a variety of file types using a single reproducible document known as a Quarto document (.qmd file). While RStudio is capable of rendering documents into numerous file types, we will generally focus on HTML, Word, and PDF in this class.

   You will be expected to turn in a .qmd file and .pdf file for all assignments in this course - any other file type will not be graded. In order to render to a .pdf file, RStudio relies upon a LaTeX distribution. There are a variety of LaTeX distributions available, but I

recommend using `tinytex`, as it is made specifically for use with R. To install `tinytex`, run the following code chunk:

```
install.packages("tinytex")
tinytex::install_tinytex()
```

> **!** STOP
>
> Before continuing, click the render button at the top of the *editor* panel and ensure that you are able to render this document to a .pdf file.

Combining languages

3. When creating .qmd files, you can use a variety of different languages to produce desired effects. For example, you may use Markdown to typeset the .qmd file (in fact, the `[]()` shortcut used to hyperlink a website *is* Markdown). You are also able to use other languages to typeset the document, which depend on what type of document you are rendering. For example, when rendering to .pdf, you may use LaTeX commands; when rendering to .html, you may use HTML code.

> **!** STOP
>
> Copy and paste the entire YAML header (which begins with `---` as the first line of this document, and ends with `---` on line 23 of this document) into a separate R script (under the file tab) so that you have it for later.

Render this document first as a .pdf file by clicking the render button. Which languages rendered as boldface?

- Markdown: **The quick brown fox jumps over the lazy dog.**
- LaTeX: **The quick brown fox jumps over the lazy dog.**
- HTML: The quick brown fox jumps over the lazy dog.

Next, render the document as a .html file by replacing

```
format:
 pdf:
 fontsize: 12pt
 geometry:
   - inner=1.5cm
   - outer=1.5cm
   - top=2.5cm
   - bottom=2.5cm
 include-in-header:
```

```
  - text: |
      \addtokomafont{disposition}{\rmfamily}
      \usepackage{fancyhdr, lastpage, framed, caption}
      \captionsetup[figure]{labelformat=empty}
      \pagestyle{fancyplain}
      \fancyhf{}
      \lhead{\fancyplain{}{STAT 0116: Introduction to Statistical Science}}
      \rhead{\fancyplain{}{Stratton - Day 2}}
      \fancyfoot[RO, LE] {page \thepage\ of \pageref{LastPage}}
      \thispagestyle{plain}
```

with

```
format: html
```

in the YAML header and clicking render. Which languages rendered as boldface?

Learning LaTeX, Markdown, or HTML is not the focus of this class. However, you will be exposed to a fair amount of each language, and each of these languages are marketable with employers. I encourage you to play around with typesetting documents in RStudio!

> **!** STOP
>
> Replace the updated YAML header with the original code so that this document renders to .pdf. Ensure that the document renders to .pdf before continuing. Note that the YAML header is sensitive to indentation. If the document will not render to .pdf, please notify me.

Regression review

HERE

```
library(tidyverse) # for data manipulation
library(MASS) # for data
data(geyser) # get the data
head(geyser) # first six rows
```

```
  waiting duration
1      80 4.016667
2      71 2.150000
3      57 4.000000
4      80 4.000000
5      75 4.000000
6      77 2.000000
```