Please answer the following questions, as this will give us a better understanding of the book:

1. Proposed title: **R for Non-Programmers: A Guide for Social Scientists (R4NP)**
2. Authors: **Daniel Dauber**
3. What is the subject and academic level?

**Quantitative Research Methods, suitable for any degree level, i.e. Undergraduate, Postgraduate (Masters) and PhD. It is an introduction to the R programming language within the context of Social Sciences research projects.**

1. What is the audience; will it be used for teaching, research, or both?

**The book is mainly intended to be used for teaching purposes and as a reference book for research when trying to perform fundamental statistical techniques commonly encountered in the Social Sciences.**

1. What are the three most closely related titles? Briefly describe how your book fits into the existing literature.

The following publications are likely the closest to R4NP:

* Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for Data Science*. O'Reilly Media, Inc.
* Kennedy, R., & Waggoner, P. D. (2021). *Introduction to R for Social Scientists: A Tidy Programming Approach*. CRC Press.
* Field, A., Miles, J., & Field, Z. (2017). Discovering statistics using R. Sage Publications.

Wickham et al.'s (2023) influential work, considered a cornerstone among data scientists, caters to a more advanced audience, emphasising R programming for those well-versed in the realm of programming languages. In contrast, R4NP is tailored to absolute novices. Many current readers of R4NP have no background in programming languages. In addition, R4NP provides a valuable complement by extending coverage to statistical techniques often found in other research disciplines, especially the Social Sciences. As such, R4NP helps bridge the gap for readers with varying levels of programming proficiency and is particularly useful to scholars and students in the Social Sciences.

Kennedy & Waggoner's (2021) publication aligns closely with R4NP, emphasising the intricacies of R programming while also delving into essential statistical techniques used in the Social Sciences. Although there is some overlap in content, the distinctive feature of their work lies in the coverage of statistical challenges faced by researchers. For example, data wrangling and sources of bias, while briefly covered by Kennedy & Waggoner (2021), are common challenges found when working with complicated datasets and are received very prominent sections in R4NP. From a pedagogical perspective, R4NP also introduces interactive elements, including chapter exercises in the accompanying R package, facilitating readers internalising this new programming language and statistical techniques. This interactive approach, particularly beneficial for novices, enhances the overall learning experience and distinguishes it as a valuable resource. The accompanying R package also includes all datasets covered in the book, which allows all readers to copy and paste code blocks and try them in their own version of R on their computer.

Field et al.'s (2017) book, concentrating on diverse statistical techniques with less emphasis on programming, complements R4NP's comprehensive approach. While lacking interactive features, it caters to readers interested in statistical methods, utilising base R predominantly and providing an alternative perspective for slightly more experienced programmers.

Lastly, R4NP uniquely addresses a particular gap in literature not covered by any of the aforementioned books: Mixed-methods research. R4NP devotes an entire main chapter to it, recognising this increasing trend in the Social Sciences. R4NP stands out for its emphasis on this essential aspect, showcasing the strength of R programming in accommodating diverse research methodologies.

1. Please provide an outline of chapters in the book, with sub-headings if possible.

The current version of the book has the following chapters and sub-headings:

**Welcome**

**About the author**

**Acknowledgements**

I SETTING THE STAGE

1. Readme. before you get started
   1. A starting point and reference book
   2. Download the companion R package
   3. A 'tidyverse' approach with some basic R
   4. Understanding the formatting of this book
2. Why learn a programming language as a non-programmer?
   1. Learning new tools to analyse your data is always essential
   2. Programming languages enhance your conceptual thinking
   3. Programming languages allow you to look at your data from a different angle
   4. Learning any programming language will help you learn other programming languages

II GETTING STARTED

1. Setting up R and RStudio
   1. Installing R
   2. Installing RStudio
   3. When you first start RStudio
   4. Updating R and RStudio: Living at the pulse of innovation
   5. RStudio Cloud
2. The RStudio Interface
   1. The Console window
   2. The Source window
   3. The Environment / History / Connections / Tutorial window
   4. The Files / Plots / Packages / Help / Viewer window
   5. Customise your user interface
3. R Basics: The very fundamentals
   1. Basic computations in R
   2. Assigning values to objects: '<-'
   3. Functions
   4. R packages
      1. Installing packages using `install.packages()`
      2. Installing packages via RStudio's package pane
      3. Installing all necessary \*R\* packages for this book
      4. Using R Packages
   5. Coding etiquette
   6. Exercises
4. Starting your R projects
   1. Creating an R project file
   2. Organising your projects
   3. Creating an R Script
   4. Using R Markdown
5. Data Wrangling
   1. Import your Data
      1. Import data from the \*Files\* pane
      2. Importing data from the \*Environment\* pane
      3. Importing data using functions directly
   2. Inspecting your data
   3. Cleaning your column names: Call the 'janitor'
   4. Data types: What are they and how can you change them
   5. Handling factors
      1. Recoding factors
      2. Reordering factor levels
   6. Handling dates, times and durations
      1. Converting dates and times for analysis
      2. Computing durations with dates, times and date-time variables
   7. Dealing with missing data
      1. Mapping missing data
      2. Identifying patterns of missing data
         1. Missing completely at random (MCAR)
         2. Missing at random (MAR)
         3. Missing not at random (MAR)
      3. Replacing or removing missing data
      4. Two more methods of replacing missing data you hardly ever need
      5. Main takeaways regarding dealing with missing data
   8. Latent constructs and their reliability
      1. Computing latent variables
      2. Checking internal consistency of items measuring a latent variable
      3. Confirmatory factor analysis
      4. Reversing items with opposite meaning
         1. Reversing items with 'mutate()
         2. Reversing items with 'recode()\*
         3. Reversing items with the 'psych' package
         4. A final remark about reversing item scores
   9. Once you finished with data wrangling
   10. Exercises

III ANALYSIS

1. Descriptive Statistics
   1. Plotting in \*R\* with 'ggplot2\*
   2. Frequencies and relative frequencies: Counting categorical variables
   3. Central tendency measures: Mean, Median, Mode
      1. Mean
      2. Median
      3. Mode
   4. Indicators and visualisations to examine the spread of data
      1. Boxplot: So much information in just one box
      2. Histogram: Do not mistake it as a bar plot
      3. Density plots: Your smooth histograms
      4. Violin plot: Your smooth boxplot
      5. QQ plot: A 'cute' plot to check for normality in your data
      6. Standard deviation: Your average deviation from the mean
   5. Packages to compute descriptive statistics
      1. The 'psych' package for descriptive statistics
      2. The 'skimr' package for descriptive statistics
   6. Exercises
2. Sources of Bias: Outliers, Normality and other 'Conundrums'
   1. Linearity and additivity
   2. Independence
   3. Normality
   4. Homogeneity of variance (homoscedasticity)
   5. Outliers and how to deal with them
      1. Detecting outliers using the standard deviation
      2. Detecting outliers using the interquartile range (IQR)
      3. Removing or replacing outliers
      4. Concluding remarks about outliers
   6. Exercises
3. Correlations
   1. Plotting correlations
   2. Computing correlations
   3. Significance: A way to help you judge your findings
   4. Limitations of correlations
      1. Correlations are not causal relationships
      2. Correlations can be spurious
      3. Simpson's Paradox: When correlations betray you
   5. Exercises
4. Power: You either have it or you don't
   1. Ingredients to achieve the power you deserve
   2. Computing power
   3. Plotting power
   4. Concluding remarks about power analysis
   5. Exercises
5. Comparing Groups
   1. Comparability: Apples vs Oranges
   2. Comparing two groups
      1. Two unpaired groups
      2. Two paired groups
   3. Comparing more than two groups
      1. Multiple unpaired groups
      2. Multiple paired groups
         1. Testing the assumption of sphericity
         2. Visualising and computing multiple paired group comparisons
   4. Comparing groups based on factors: Contingency tables
      1. Unpaired groups of categorical variables
      2. Paired groups of categorical variables
      3. A final remark about comparing groups based on categorical data
   5. Reviewing group comparisons
   6. Exercises
6. Regression: Creating Models to Predict Future Observations
   1. Single linear regression
      1. Fitting a regression model by hand, i.e. trial and error
      2. Fitting a regression model computationally
   2. Multiple regression
      1. Outliers in multiple regressions
         1. Outliers in the dependent variable
         2. Outliers in independent variables
         3. Outlier detection considering the entire model: Global measures
         4. Outlier detection considering the entire model: Specific measures
         5. Reviewing the outliers
      2. Standardised 'beta' vs. unstandardised 'beta'
      3. Multicollinearity: The dilemma of highly correlated independent va...
   3. Hierarchical regression
      1. Regressions with control variables
      2. Moderated regression
      3. Centering predictors: Making 'beta's' more interpretable
   4. Other regression models: Alternatives to OLS
   5. Exercises
7. Mixed-Methods Research: Analysing Qualitative Data in R
   1. Mixed-methods research: Analysing qualitative data in \*R\*
   2. The tidy process of working with textual data
   3. Stop words: Removing noise in the data
   4. N-grams: Exploring correlations of words
   5. Exploring more mixed-methods research approaches in \*R\*
   6. Exercises

IV MOVING BEYOND

1. Where to go from here: The next steps in your R journey
   1. GitHub: A Gateway to even more ingenious \*R\* packages
   2. Books to read and expand your knowledge
   3. Engage in regular online readings about \*R\*
   4. Join the Twitter community and hone your skills

**Epilogue**

**APPENDIX** (which will include additional case studies)