Grading Guidelines for STV2020

The exam form in STV2020 is a term paper which the students have worked on throughout the semester and which they have been offered the chance to receive feedback on during the semester. The term paper needs to be submitted both as a Quarto file (.qmd) and as a rendered document (.html). Please consider both he Quarto and output document when grading the term paper: students have been advised that the output document should look "polished" and need to display all figures and tables discussed on the text. The code in the Quarto file may, however, contain code not displayed/discussed in the output document.

The term paper should be evaluated based on the "data science" skills the students demonstrate in the paper, *i.e.* their ability to collect, manipulate, and visualize data using R and to document their workflow and results in Quarto. Please note that this is not a statistics course. Students are not expected to fit or interpret statistical models. Please also note that this is not a substantive social science course. Situating the term paper in a social science literature is not important and students should not be penalized for picking research questions without a clear social science relevance. It is, however, a technical course: writing good and efficient R code, demonstrating familiarity with different data science techniques, and to produce beautiful and instructive visualizations and output documents. The following questions should be used for assessing the term papers:

- Does the term paper demonstrate a good command of key techniques covered in the course, such as data wrangling, scraping static websites, extracting information from text, interacting with APIs to retrieve data, spatial data, and data visualization? It is not necessary for a term paper to use all these techniques and it is not desirable for students to "force" an application that is not really useful for their project. However, good term papers should demonstrate an ability to apply different techniques. If students go beyond the techniques covered in the course to solve a problem that arise in their project (e.g. figure out how to scrape a dynamic website), this is something to be rewarded.
- Does the code run? Is it well documented/commented? Are there errors that the students did not find/correct?
- Is the code efficient? Students have been told to avoid copy/paste and to use user-defined functions, iteration, helper verbs in dplyr, etc. to make their code readable.

- Are the visualizations helpful in conveying a message/illustrating the data? Are they aesthetically pleasing? Students have been told to rather make a few illustrations that are suitable for conveying a message than to just produce every visualization imaginable.
- Did the Quarto file render correctly? It is important that the figures, tables, bibliographies, etc. are actually displayed in the output also when the document is exported. Good term papers will also use cross-referencing, reference management etc.

The best term papers will score high on all the above indicators. However, please keep in mind that most students have very limited experience with R and programming prior to this course. A lot of time has been spent trouble shooting errors and just making the code run!