|  |  |
| --- | --- |
| **Sections** | **Standard** |
| **Introduction** | - Provide an informative introduction that explains the problem statement you’re answering. Why should I be interested in reading your report?  - Provide a short explanation of how you’re addressing this problem statement, what data are you using, what methods are you using?  - Discuss your approach and analytic techniques and why it will address this problem.  - Explain how your analysis will help the reader of the analysis |
| **Packages Required** | - All packages are loaded at the beginning of the code  - Messages and warnings from loading these packages are suppressed.  - Explanation is provided for each package in the comments. Don’t assume I know why you loaded it! |
| **Data Preparation** | - Source of the data is cited and hyperlinked, if possible. Source data is thoroughly explained (who collected it, why was it collected, when was it collected)  - Data import and cleaning steps, if necessary, are explained in the text and follow a logical process. Tell me why you’re doing what you’re doing.  - Show me the data set in the most condensed form possible. (Do NOT print out a data frame with 200+ rows)  - Provide a summary of the variables you’re interested in. Don’t just do str() or summary(), but provide me something like an explanation with a table or a nice summary paragraph. |
| **Exploratory Data Analysis** | - Uncover new information in the data that isn’t immediately apparent (such as: don’t just plot the data as is, but filter or group the data in different ways, create new variables, or join different data frames together to get new summaries)  - Provide your findings in plots or tables. Display findings in different ways.  - Graphs all have purpose and help tell your story. Make sure you have graphs that illustrate your primary points and that they are all appropriately formatted (plot and axis titles, legends if needed, appropriate plot types used)  - Use tables effectively to perform comparisons, highlighting important features and findings.  - Your analysis should be thoroughly, yet concisely, explained. Make sure your insights and discoveries are highlighted, and easy to see and understand. |
| **Summary** | - Summarize the problem statement, and how you addressed this problem statement  - Summarize the interesting insights your analysis gave  - Summarize the implications of your findings  - Discuss the limitations of your findings and how you or someone else could improve upon it |
| **Formatting and Other Requirements** | - Code is well commented and proper coding style is followed.  - Coding is systematic – complicated problems broken down into simpler sub-problems, the code is efficient, correct, and minimal.  - Achievement, mastery, cleverness, creativity: Tools and techniques from the course are applied very competently and, perhaps, somewhat creatively.  - Rmd fully executes without any errors and html/pdf produced matches the html/pdf report submitted by student.  - OR Word document contains relevant code snippets where necessary and is formatted clearly, and is provided along with a cleanly commented code script that was used to produce the results. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Points:** | **1** | **2** | **3** | **4** | **5** |
| **Introduction – Problem Statement** | Barely present intro statement, doesn’t address the problem or the data | Introduction that discusses the problem or the data, but not both. | Introduction that discusses the problem, refers to the data | Introduction that describes the problem directly, outlines what data you’re using to address it | Interesting, clear, engaging, informative introduction that describes the problem directly, clearly outlines what data you’re using to address it |
| **Introduction – Approach and Explanation** | Introduction does not address the tools used |  | Introduction addresses and introduces the tools used |  | The introduction gives a clear preview of the rest of the report, and explains why the tools chosen are appropriate |
| **Data Preparation: Source and cleaning** | Source data is presented by a brief mention. Source of data is not discussed. Data cleaning that should be done is not done. | Source of the data is cited and hyperlinked, if possible. Source of data is explained poorly if at all, data cleaning is not explained or done well. | Source of the data is cited and hyperlinked, if possible. Source of data is explained. Data import or cleaning occurs without explanation. | Source of the data is cited and hyperlinked, if possible. Source of data is explained.  Data import and cleaning steps, if necessary, are explained but are haphazard or ad hoc. | Source of the data is cited and hyperlinked, if possible. Source of data is thoroughly explained Data import and cleaning steps, if necessary, are explained and follow a clear logical process. The report explains why you’re doing what you’re doing. |
| **Data: Display and summary** | Minimal or no effort is made to summarize the data’s contents. | The data is shown The variables are listed using nothing more than str(), summary(),k or glimpse() | The data is displayed. The variables of interest are listed. | The data is displayed in a condensed fashion. The variables of interest are summarized and explained. | The data is displayed and summarized cleverly and well. The variables of interest are summarized and explained skillfully. |
| **Exploratory Data Analysis: Discovery** | The report explores the data simply, at the surface level, and does not try for more. | The report makes an basic but unsuccessful attempt at exploring the data | The report makes an basic yet successful attempt at exploring the data in a new way. | The report uncovers new information in the data that isn’t immediately apparent | The report successfully finds interesting, intriguing, or noteworthy relationships in the data. |
| **EDA: Plots, Graphs, and Tables** | Graphs and tables might be used incorrectly, graphs are hard to read or understand. | Findings are displayed using the correct type of plot or tables. Plots have a few flaws or errors, and they detract from the message, hiding the main point of the plots. | Findings are displayed using the correct type of plot or tables. | Findings are displayed using the correct type of plot or tables. Plots might have a few flaws or errors, but they don’t detract too much from the message. | Plots and tables are all clear, easy to understand, look nice, and play a role in the analysis. Plots have titles, axis labels are all correct, appropriately sized and formatted well in the scope of the paper. |
| **EDA: Explain** | No explanations are given for the steps taken. Insights and discoveries are not mentioned. |  | The analysis is clumsily explained. Insights and discoveries are mentioned, but it is hard to find them or they are vague and unclear. |  | Your analysis should be thoroughly, yet concisely, explained. Make sure your insights and discoveries are highlighted, and easy to see and understand. |
| **Summary: The problem and the insights** | There is little to no attempt to summarize the problems and the findings. | There is a summary, but it is vague, brief, or unclear. | There is a summary of the findings, the insights about the questions are present and restated. | The summary of the findings is pretty good, and the insights about the questions are present and synthesized. | There is a concise, clear, thorough summary of the problem and the findings, and the insights from the findings are detailed and exhaustive. |
| **Summary: The implications and the limitations** | A very basic summary is stated, if at all. There is no discussion of the limitations of the analysis. | A very basic summary is stated. There is a brief mention of the limitations of the analysis. | Facts about the findings are stated, and the limitations of your analysis are mentioned. | Conclusions about and implications of the findings are drawn, and the limitations of your analysis are discussed. | Solid, logical conclusions about and implications of the findings are drawn, and the limitations of your analysis are discussed in detail. |
| **Formatting and Other Requirements** | The report is chaotic to read and follow, didn’t compile/turned in as an .Rmd, no comments/whitespace/etc. in the code. | The report has solid overall structure, but has a large number of typos or formatting issues, or the code is difficult to follow/understand. | The report has the occasional typo, the code is mostly easy to follow. Tools and techniques from the course are applied. | The report looks nice, easy to follow code and good sectioning and layout. Tools and techniques from the course are applied competently. | The report is professional, polished, with easy to follow code and clear sectioning and layout. Tools and techniques from the course are applied very competently and creatively. |