Non-Conformance Report

For ISYS2001 Business Programming
Assignment Two Notebook Report/Code
Semester One 2022

Prepared By:
Tom Sargent - 20594567

Table of Contents

Non-conformance Report Entry 1

Date: 10/05/2022	Entry Number: 1
The Problem	After working on the assignment and processing the data in Google Collab I have noticed that some of the date formats do not conform to the standard Pandas Date/Time layout. On discovery I have found that pandas has a unique date-time format that the module is able to read and analyse. Example: Some of the years do not conform to datetime values 0001-02-04 (Where year is not valid), possibly instead of 2001-02-04 or outside of the normal range.
Severity and Impact	This will affect the report, as without finding a solution, the graphs cannot be generated. I need to remove the timestamp from the datetime format, however, I need to find a way to filter or ignore the entries that cannot be accepted.
How it Occurred	Occurred when testing the Value.counts function and found many dates that don't conform. I also found that when trying to convert date time column to Pandas Datetime to remove the time so we can plot date, that this issue occurred. The issue itself is likely to do with the dataset, potentially relating to database compatibility issues at the time of the email recording in the early 2000's
How to prevent it from happening	I can prevent this by researching stack overflow how to filter/remove dates that don't conform to datetime, possibly even ignoring exceptions.
Plan or time estimate to fix	1-2 Days after some further research and analysis. **UPDATE:** This issue is now fixed – I have worked out how to filter by date range, so the dataframe ignores all of the dates that do not fall within a the given period of the Enron Dataset timeline. After this filter, the value.counts function and the plots will generate successfully.

Non-conformance Report Entry 2

Date: 24/05/2022	Entry Number: 2
The Problem	After almost completing the notebook, I have noticed that the report lacks exception handling. Although I have tested running the notebook many times, with no errors occurring. I know it's important to include proper exception handling in the event of an error, as it is both good practices, and essential to the smooth operation of a piece of code. I know how to do basic try, except exception handling however I want to try and find more complex ways to demonstrate my understanding of the concept, perhaps designing for specific types of errors, or providing feedback.
Severity and Impact	This will affect the report, as without proper exception handling, in the event of an error, the user will experience an interruption.
How it Occurred	This issue occurred when I realised that the report lacks well-constructed exception handling and it is a requirement of both the marking guide and good programming practice.
How to prevent it from happening	We can prevent this issue by researching, practicing and including more exception handling within the report.
Plan or time estimate to fix	I plan to try and resolve this before the submission date and try to include some more exception handling. If I am unable to do so before this date, I hope to work on this more personally into the future, to increase my coding skills.

Non-conformance Report Entry 3

Date: 25/05/2022	Entry Number: 3
The Problem	I have noticed that although my graphs generate, they lack data highlighting that I would have liked to achieve, as I know how to do in programs such as Excel and Power BI. It would be impactful to be able to highlight data points on graphs to show the reader areas of importance, which can then be discussed within the report itself.
Severity and Impact	This has a low impact to the completion of the report, however, It is a feature that I would like to learn how to use and add to the report.
How it Occurred	This issued occurred when I was reading the report, and realised it needed further analysis to strengthen the report's findings from the dataset. Highlighting and analysis of graphs is crucial, so if I am able to display an understanding of the data clearly, I will have a well written report.
How to prevent it from happening	I can continue to learn Matplot lib coding examples to increase my understanding of the program from both websites and in class examples that we have completed previously.
Plan or time estimate to fix	I plan to try and resolve this before the submission date and try to include some more exception handling. If I an unable to do so before this date, I hope to work on this more personally into the future, to increase my coding skills.

Non-conformance Report Entry 4

Date: 25/05/2022	Entry Number: 4

The Problem	Another issue that I am facing is the generation of plots showing at different sizes. It appears that when I run all on the notebook, the plots may generate at a different size/aspect ratio randomly. This is not ideal as it does not allow for consistency in the report, and can make a regular graph look very disproportionate.
Severity and Impact	This likely has a medium to high impact, as it disrupts the clean flow of the business report, when graph sizes are inconsistent.
How it Occurred	This occurred when I wrote the code for the matplotlib graphs, and in doing so realised that they generate at different sizes each time.
How to prevent it from happening	I have done some investigation on stack overflow and the matplot lib website itself and so far I have found the following. - You can define each graph size or you can define the maximum parameters for the matplotlib module in which all graphs will generate. - Sometimes graphs generate differently due to data size.
Plan or time estimate to fix	I estimate to try and find a solution to this before the due date, or at least understand a compromise and how I have gone wrong in being unable to achieve this. **UPDATE:** 28/5/22 – After adjusting some settings and the rcPARAMS it appears I have some more consistency in graph size. I also found out how to make the graph dpi higher, making it easier to read and understand.