You have landed an internship providing analytics for General Mills (GM), a major cereal producer. They are concerned about the effectiveness of in-store promotions and advertisements (the cost of the programs is complex and not directly attributable to promotions or advertisements). Your manager, Quian Xu, needs to know how GM is doing generally relative to their primary competitors and specific strategies for GM cereals. If there is a promotion it indicates that price is discounted by at least 5% (note the price listed is what is paid at the check-out). GM has provided you with a representative sample of weekly (lowest week number is first week of the year) sales from several stores that has been anonymized so there are no privacy concerns for customers or stores. Analyze the data, determine two *significant* (practical and statistical to be discussed in week 5) findings related to the effectiveness of advertising and promotions, create a professional quality visual for each finding, and provide a brief write up of your findings and recommendations in a memo using best practices. It may be beneficial to separate out the producer from the brand to determine if there are producer name or brand effects on sales, promotions, ads, or prices.

Data: mtp_product_data.csv and mtp_store_data.csv, the variables are listed below

store_key: store number	**UPC**: unique product number	**week**: week of sale
units: number of cereal packages sold	**brand**: producer and brand	**promo**: in store promotion 0/1 is no/yes
price: price per package	**flavor**: cereal flavor group	**ad**: none A – medium B – small
volume: cereal package size	**package**: type of cereal container	

Memorandum Format: 2-page maximum including graphs and/or tables with appropriate memo header for a non-technical audience (HTML doesn't have page breaks so use your judgement to determine page length). Your memorandum should include the following:

1) An introductory paragraph giving background, brief description of the data, and brief description of your findings. This paragraph is meant to engage your audience so it should make the analysis compelling—*why is the problem important to Quian and GM.*
2) A paragraph describing each of your two most important findings, each supported by a visualization (and possibly a small table) that are explained in non-technical terms.
3) Conclusions and suggested course of action in the form of a recommendation.

Technical Appendix

4) The technical appendix contains your base, detailed, and statistical EDA, and the visuals to load to your memo. It is important that your work is clearly organized (use a table of contents (TOC)), easy to understand (code clearly documented), and reproducible (all in R, no cutting and pasting). Title each step of your analysis so it appears in the TOC. The files must run from the original data sets, so all data wrangling must be contained in the RMD files. The technical appendix explains what you are doing and why, technical jargon is acceptable because the audience is expected to have expertise. Write out what you observe and questions that arise after each step during your base EDA – these are the data comments/questions that you will examine in your detailed EDA. Think of each code chunk, output, and your written observations as a page in your notebook. Test your findings statistically (week 5), create your professional visuals, and save them externally.

Submissions: Submit the html and RMD files to Canvas—title files "LastName_LastName_TA.RMD", "LastName_LastName_TA.html, "LastName_LastName_memo.RMD, and "LastName_LastName_memo.html for each person in your group and include both of your names in the TA and memo header. I will knit the file to make sure it runs properly.

Remember, all writing in the memo must be clear, concise and accurate – no repetition or excessive use of adverbs and adjectives. No fluff or filler, every sentence and word should have purpose. **Quian does not have time to guess at what you have done, and she is not a data analyst, so the memo must be written clearly in plain English with NO technical jargon.** However, the technical appendix can include jargon, but must be clearly documented so readers do not have to guess what you have done and why.

Category	Score	**Comments**
Introductory Paragraph – (3) Compelling (1), brief finding (1), data (1)		
Finding 1 – (8) Visual (4), description (2), clear (1), exceed (1)		
Finding 2 – (8) Visual (4), description (2), clear (1), exceed (1)		
Conclusion – (3) Summary (1), impact (1), contact (1)		
Technical Appendix – (24) Clear organization (2), documentation of code (3), data questions (4), join data (2), basic EDA (4), detailed EDA (4), stat EDA (3), exceed expectations (2)		
Mechanics of memo and TA – (4) Grammatically (1) and typographically (1) correct, effective memo design (1), no cut & paste (1)		
Total	**/50**	

What does "clear" mean? When you are writing to a lay audience you need to assume they have only a limited background in quantitative methods. You can generally assume that people know what an average is and understand that variance or standard deviation relates to variation of the data, but not much more. They may or may not know what the median is and why it might be better than the mean at representing data. You would likely need to explain what a box plot or histogram illustrates. When trying to determine if terminology or a visualization is "clear" or not, think about saying or showing it to a friend without a quantitative background. Would they know what the terminology means, or would you need to provide more background for them to understand it? In fact, a good way to determine if your memo is written appropriately is to have a non-technical person read your memo. Clear also means your writing is professional – no slang, typos, and good grammar. The Seattle University writing center is a good group to work with.

Point Allocation

100% – Meets all expectations **and exceeds** some for engagement and exploration
87.5% – Meets all expectations for engagement and exploration
75% – Meets most expectations for engagement and exploration
50% – Meets some expectations for engagement and exploration
25% – Does not meet any expectations for length, engagement and exploration