You have recently begun an internship for an online office supply store, Office Mate. They have been in business since 2012 and have done well. However, they are beginning to face more competition from Amazon and several other online retailers, so they need to determine how to become more profitable. Your manager, Quian Xu, has come to you and a co-worker to complete an analysis and come up with two recommendations to improve profitability. The specifics of the task are outlined below. Quian will be sharing your findings with upper level management and the board of directors, so it is critical that your writing and graphs are clear to people without a technical background. Analyze the data, determine two *significant* findings (practical and statistical to be discussed in week 5) related to profitability, create a professional quality visual(s) for each finding, and provide a brief write up of your findings and recommendations in a memo using best practices. You may work in teams of two. Use the data "mtp_off_mate.csv", which includes the following variables:

1. Order ID: Each order ID is unique, but may contain several products
2. Order Date: Date the order was placed
3. Ship Date: Date order was shipped
4. Ship Mode: Shipping method
5. Customer ID: Each customer has a unique ID, but can make more than one order
6. City, State, Postal Code: For the location of each sale in continental US
7. Region: Sales region in the United States (Central, East, South, West)
8. Product ID: Product ID code
9. Product Name: Name of product sold
10. Segment: Customer segment (Consumer, Corporate, Home Office)
11. Category: Type of product sold (Furniture, Office Supply, Technology)
12. Sub-Category: Detailed type of product sold
13. Revenue = Price * Quantity (though the dataset does not include Price)
14. Quantity = number of units of the specific product sold for that order
15. Discount: Percent discount from full Price. Large discounts are used to clear inventory.
16. Profit = Revenue – Cost (not including shipping or tax, both passed directly to customer)

Each observation is for the sale of a specific product from an order that may contain multiple products. The Order ID variable links products from the same order. The customer ID identifies unique customers that may have multiple orders

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 Office Mate*.
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, and possibly other minor findings.

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 using the following method: "LastName_LastName_TA.RMD", "LastName_LastName_TA.html, "LastName_LastName_memo.RMD, and "LastName_LastName_memo.html.
Also include both of your names in the TA and memo headers. I will knit the file to make sure they run properly, so be careful to make sure it all works from the original data.

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), data (1), brief finding (1)		
Finding 1 – (8) Visual quality (4), descrip (2), measure (1), exceed (1)		
Finding 2 – (8) Visual quality (4), descrip (2), measure (1), exceed (1)		
Conclusion – (3) Summary (1), recommend (1), additional (1)		
Technical Appendix – (24) Clear organization (2), documentation of code (3), data comments/questions (4), data wrangling (2), base EDA (4), detailed EDA (3), stat EDA (3), exceed expectations (3)		
Mechanics of memo and TA – (4) Grammatically (1) and typographically (1) correct, effective memo design (1), no cut & paste (1)		
Total	**/50**	

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