MARK5828 Advertising Analytics Individual Research Project CRITERIA

<1> Interim Report (Total Marks = 5) You need to submit a report for your interim. Based on your report, tutors and your group members will give you feedback.	**Marks**
Research Question? 1. Topics which MARK5828 Advertising Analytics talked about. For example, a. Effective picture or video ad for consumer engagement in social media? (W1, W2) b. Measuring the effect of a commercial or non-profit ad campaign using consumer search or mass or social media? (W3, W4) c. Other topics using image/video are also possible. 2. Types of research questions a. Discovering several X factors i. Exploring a large set of X's ii. What's the importance order? iii. (e.g.) video content (actor, colour, words) to increase the number of like? b. Causal effect i. Does a particular X affect Y? ii. Deep investigation to show whether X indeed affects Y iii. (e.g) Is it good for minority founder (e.g. African American) to create their own Video ad in Kickstarter?	1
Data in Moodle 1. Saatchi Art pricing ● Consists of 3 datasets & 3 photography datasets ● Please choose 1 dataset to work with 2. Kaggle Cartoon 3. Sydney Instagram 4. UNICEF Instagram 5. Climate Change Instagram 6. Kickstarter video for African American Female Founder 7. Kaggle Movie TMDB Box Office Prediction (movie summary, actor..) 8. If you have data for your research, other projects, and so on, you can use it. 9. We also encourage you to search for picture/video data. If you search data, please make sure whether the data has a dependent variable (e.g. number of likes and so on) a. Google Data Set Search https://toolbox.google.com/datasetsearch b. Kaggle **c.** AWS **d.** UCI ML repository **e.** Social media datasets **f.** Youtube8M **g.** Image data for computer vision **h.** BigML **i.** Image and Video dataset 10. You could collect data using API. In this case, you should not use this collected data for commercial purposes. If so, you will get 0 for your project. a. Python Beautifulsoup (Web Scraping) b. Facebook Graph API (Extracting information from Facebook) c. Twitter streaming API (Extracting Information from Twitter)	

11. Human coding - You can simply visit your own or particular companies' social media (e.g. Facebook, Pinterest, Instagram) and then make variables for yourself. - e.g.) Whether a particular object exists in the picture or not. - e.g.) Human perception/evaluation of visual content - Don't forget to collect social media engagement outcome (e.g. likes, pins, shares, comments, re-Tweets). **NOTE 1: You are required to use a DIFFERENT DATASET from your group project data. If you use the same data, you will get 10 PENALTY POINTS from your final report mark.** **NOTE 2: Unlike the Group Project, you do not need to email your tutor the dataset you will be using.**	
Data Collection: Relevant data? 1. What is your data? a. What does each column represent? A table is highly recommended for this. (eg, The *likes* columns represent the amount of user likes on a YouTube video; *popularity* is *views* x *rating,* etc) 2. How did you collect it? If you're using a Moodle dataset, please still inform the marker where the data actually came from, and don't say 'I am using the Sydney Instagram dataset from Moodle' (eg, I used the Twitter API to extract every tweet from Obama from 2012 to 2018) 3. Do you have relevant data to answer your research questions?	1
Data Exploration and Cleaning 1. Please make at least 5 plots using Power BI or Python. 2. What do you learn from your plots? 3. Based on the plots, are you going to revise your hypothesis? Or make a new hypothesis? 4. Did you clean your data? Please state your data cleaning process clearly. A flow chart may be useful here. You may want to consider if you need to: a. Dropping Outliers or categories with only a few observations b. Removing columns that are not suited to your goal/hypothesis c. Dealing with missing values d. Adding new columns 5. From the plots, do you feel that you need to analyze a subset of data? For example, we split UNICEF video into two parts in Week 2. But, this is optional.	1
Initial Result 1. Run regression (use VIF) and interpret your result. a. Use ALL the variables which you think are important to your goal. b. Ensure that no columns have a Variance Inflation Factor higher than your defined threshold. 2. Does this result support your hypothesis?	1
What is your plan for your final presentation and report? 1. Do you need to revise your project goal/research questions (hypothesis)? 2. Do you need to collect additional data or clean data more? 3. What additional analysis are you planning? 4. **What is your schedule by week**?	1

	Task
W6	
W7	
W8	
W9	
W10	

Submission

- You are required to submit:
  - A **.pdf** of your report. At the beginning of your report, you are required to include the names and zIDs of all your group members.
  - A **.ipynb** containing all relevant code to get the results in your report.
  - A **.csv** file containing the dataset you run regression on. (use df.to_csv("outfile.csv")).
- Also, make several copies of your report for your group members. They will give you their feedback during the tutorial time.

Deadlines

- **Due: Thursday 10pm Week 8 (11th April)**
  - Please do your best to meet this deadline so that your tutor can have enough time to think about comments on your report.
  - Cutoff: Friday 5pm Week 8 (12nd April) (Late penalty: -1 mark)
- You will NOT be able to make a submission after the cutoff date.

<2> 3 min Final Individual Research Presentation + 2 min Q&A (Total Marks = 5)	**Marks**
Data Visualization 1. Please make at least 5 plots using Power BI or Python. 2. Please highlight your key findings with graph or plots (Visual communication) rather than simply reporting numerical values or tables. 3. Do not show unnecessary plots.	2
Verbal Communication 1. A clear introduction/goal/hypothesis 2. A logical conclusion of their findings based on their hypothesis 3. Can the audience understand what you are trying to explain? 4. Can you answer the audience's (TAs & students) questions properly?	2
Engagement - Pitching 1. Is the pitching interesting? 2. Are the audience engaged in the presentation or not?	1
Time Management Presentations should not exceed 3 minutes.	

You will be stopped by the tutor at the 3 min 30 second mark (finish the current sentence).	
Submission - Please submit your presentation file(s) to Moodle (format is flexible). The tutor will be using these files for your actual presentation - If you want to make a presentation through other uncommon/different means, please email the tutor BEFORE the submission due date.	
Deadlines - **Due: Thursday 10pm Week 10 (25th April)** - Cutoff: Friday 5pm Week 10 (26th April) (Late penalty: -1 mark) - You will NOT be able to make a submission after the cutoff date.	

<3> Final Group Report (Total Marks = 20)	**Marks**

What is your research question?
 	
1. A topic which MARK5828 Advertising Analytics talked about. For example,
  a. Effective picture or video ad for consumer engagement in social media?
  b. Measuring the effect of a commercial or non-profit ad campaign using consumer search or mass or social media?
  c. Other topics using image/video are also possible.
 	
2. Types of research questions
  a. Discovering several X factors
  i. Exploring a large set of X's
  ii. What's the importance order?
  iii. (e.g.) video content (actor, colour, words) to increase the number of like?
  b. Causal effect
  i. Does a particular X affect Y?
  ii. Deep investigation to show whether X indeed affects Y
  iii. (e.g) Is it good for minority founder (e.g. African American) to come to their own Video ad in Kickstarter?
	2

Why is your Research Question important? Justify why your research question is important 1. Either Academic, Industry (i.e. practical or managerial), or Society (i.e. community) value is fine. 2. Cite academic papers, industry report, newspaper articles, or reliable statistics 3. Or you can show evidence from your data **Example** Younkin and Kuppuswamy (2018) found that minority founders (e.g. African-American Women) have a disadvantage in crowdfunding. To help such minority founders, I attempt to find an effective	2

Kickstarter video content strategy. Specifically, I ask whether it is good for minority founders to come in their own video advertising given that there is racial bias in the crowdfunding platform. **Reference** Peter Younkin, Venkat Kuppuswamy (2018) The Colorblind Crowd? Founder Race and Performance in Crowdfunding. Management Science 64(7):3269-3287. https://doi.org/10.1287/mnsc.2017.2774 **Academic journals** Google Scholar: https://scholar.google.com.au/ Scopus: https://www.scopus.com/ Computer vision journal Management Science: https://pubsonline.informs.org/journal/mnsc Marketing Science: https://pubsonline.informs.org/journal/mksc Information System Research: https://pubsonline.informs.org/journal/isre Journal of Marketing Research: https://www.ama.org/publications/JournalOfMarketingResearch Journal of Product Innovation Management https://onlinelibrary.wiley.com/journal/15405885 Service Science: https://pubsonline.informs.org/journal/serv You can also look at many other academic journals.	
Data Collection: Relevant data? 1. What is your data? a. What does each column represent? A table is highly recommended for this. (eg, The *likes* columns represent the amount of user likes on a YouTube video; *popularity* is *views x rating,* etc) 2. How did you collect it? If you're using a Moodle dataset, please still inform the marker where the data actually came from, and don't say 'I am using the Sydney Instagram dataset from Moodle' (eg, I used the Twitter API to extract every tweet from Obama from 2012 to 2018) 3. Do you have relevant data to answer your research questions? 4. If you don't have a relevant variable in given data, considering your research question, please find or generate additional X variables (human-coded, API, merge with online datasets, web scraping, etc) a. e.g.) Whether a particular object (e.g. brand logo) exists in the picture or not. b. e.g.) Human perception/evaluation of visual content	2
Data Exploration & Cleaning 1. Please make at least 5 plots using Power BI or Python. 2. What do you learn from your plots? 3. Based on the plots, are you going to revise your hypothesis? Or make a new hypothesis? 4. Did you clean your data? Please state your data cleaning process clearly. A flow chart may be useful here. You may want to consider if you need to: a. Dropping Outliers or categories with only a few observations b. Removing columns that are not suited to your goal/hypothesis c. Dealing with missing values d. Adding new columns 5. From the plots, do you feel that you need to analyze a subset of data? For example, we split UNICEF video into two parts in Week 2. But, this is optional.	3
Data Analysis 1. Make columns carefully for categorical variables. 2. Include necessary control X variables where applicable. 3. Run regression.	2

4. Check VIF (Variance Inflation Factor) to check for multicollinearity issues. 5. Based on VIF, choose variables and then run regression again.	
Interpretation 1. Which variables do you need to interpret to answer your research questions? 2. Do the variables affect your outcome variable (Y) significantly? 3. Does the sign of coefficient make sense? 4. Then, how much does the X variable affect your outcome variable (Y)? 5. Can you explain what your result means? Instead of showing numbers, can you produce a story about the analysis?	2
Why does particular X affect Y? 1. Can you find supporting theory from previous academic paper or industry report? 2. Or can you do further analysis to provide your own explanation? If it is hard to collect relevant data, please state at least your own hypotheses. **Example** In week 4 lecture, Junbum Kwon and Avi Goldfarb (2017) found that Ad message can impact on newspaper. While many economics research show "monetary incentive" reason, the authors show mass media's public goal (vs. profit goal) as another possible reason.	2
Future plan 1. Can you elaborate on possible future plans you might take on this project? a. If you have more time, what additional analysis do you want to do more? b. Do you want to collect more pictures or additional X variables? 2. How do you want to use your result? Making them public by writing a blog? Or Are you willing to write an academic paper?	2
Your Future Career Goal 1. What was your career goal before you take this course? 2. Now, do you have the same career goal? Or did you change? a. What is your dream job in the industry? Marketing Scientist, Media analyst, Data Scientist and so on? b. Do you consider an academic career? Professor? So do you consider a PhD in the area of Marketing, Advertising, Media, Computer Science, and so on? 3. What did you learn from the TWO projects (group and individual one)? a. Are you now more confident at analysing data? b. What were the challenges encountered during the projects? How did you overcome it? 4. Do you plan to learn more data analytics? a. What other data analytics courses are you planning to take in UNSW? (MARK5826 Product Analytics, MARK5827 Customer Analytics, Computer Science, Statistics courses?) b. What online courses are you planning to take? (Lynda.com, Datacamp, Courser, EdX, Udemy,...)	2
Report Style 1. The report has clear references to good examples/data sources and research papers. Please	1

include a reference/appendix where necessary. 2. The report is written in a clear and orderly fashion. 3. The report is highly understandable and is worthy of real-world recognition. 4. Your report should be no longer than 15 pages.	

Submission
- You are required to submit:
  - A **.pdf** of your report. At the beginning of your report, you are required to include the names and zIDs of all your group members.
  - A **.ipynb** containing all relevant code to get the results in your report.
  - A **.csv** file containing the dataset you run regression on. (use df.to_csv("outfile.csv")).
- Only ONE person in the Group is required to submit to Moodle.	

Deadlines - **Due: Tuesday 10pm Week 11 (30th April)** - Cutoff: Tuesday 10pm Week 12 (7th May) (Late penalty: -3 marks per day after due date) - You will NOT be able to make a submission after the cutoff date.	