MANG2048 (SIM) Analytics for Marketing

Project Statement

- **This group report accounts for 50% of the total marks for this module.**
- **Your group report must be typed and is limited to less than 2,000 words.**

Bike sharing has become increasingly popular across the globe and different cultures. Today, more than 550 bike sharing programmes are operating in more than 600 cities, with more than half a million bicycles. The principle of bike sharing is simple. Individuals use bicycles on an as-needed basis without the costs and responsibilities of bike ownership. It is a short-term bicycle access, which provides its users with an environmentally friendly form of public transportation. This flexible short-term use scheme targets daily mobility and allows users to access public bicycles at unattended bike stations. Bicycle reservations, pickup, and drop-off are self-service. Commonly concentrated in urban settings, bike sharing programmes also provide multiple bike station locations that enable users to pick up and return bicycles to different stations.

This project is about the Capital Bikeshare (CaBi) scheme (http://www.capitalbikeshare.com/) in Washington DC, launched in 2008 by the government of District of Columbia. In this project, we focus on the operations of CaBi in 2011 and 2012. In the Excel spreadsheet *data.xlsx*, it contains hourly rental data spanning these two years. Two types of rental data were recorded: casual renters and registered renters. A registered renter is a pre-registered member of the bike sharing program, while a casual renter pays for the use of a bike on the spot using a credit card. (Watch the YouTube link http://www.youtube.com/watch?v=qQ62cQiGFI8 to see how it works.) In addition to the time/date information (including holidays), the spreadsheet also records the weather condition in each hour (e.g., temperature, wind speed and humidity).

In early January 2013, the top management in CaBi, though enjoyed the growth of the customers, felt the need of minimising its operating costs. After hearing about how big data were used to improve the operations of some organisations from media, the top management decided to look into their own data to see if some useful business insights could be extracted. Your team was hired for this purpose. Your team was given a month to analyse the aforementioned data set collected in 2011-2012 and you were required to present your findings and insights to the top management in early February.

Part 1 (50%): Use data visualisation to explore the data as much as possible. The top management also welcomed any interesting insights that you might observe from the data. The top management also indicated that they were interested in knowing more about their customers, including the behavioural differences between casual renters and registered renters. Although you were only given limited data (not including the data by stations), the top management

wondered if the data could reveal any information about the purposes of the bike rentals by casual renters and registered renters.

Part 2 (50%): The top management would like your team to develop statistical models (e.g., hypothesis testing, ANOVA, linear regression) so that they could use them to forecast the rental demand. The top management had also indicated that they did not need an hour-by-hour forecasting model because they believed hourly demand was too volatile to forecast. Instead, if your model could forecast the daily rental demand well (for both casual and registered renters); they would be happy. The top management also said that by the time your team visited the company in February for your presentation, the data for January would have become available. You would be asked to use your model to forecast the January demand, which would then be compared with the actual demand. Assume that the January data are given in worksheet *Testing*, fill your daily demand forecast in the highlighted area in the same worksheet.

For this assignment, you are expected to answer the above questions following the advice below:

1) After identifying key variables from the data by examining patterns of interrelationship among variables, you must make corresponding justification from theory in marketing-related journals (e.g., Journal of Marketing, Journal of Marketing Research, Marketing Science, Journal of Consumer Research, Journal of Retailing, Management Science, Production and Operations Management) and use theory and the relevant literature to develop the analytic models.
2) Create workable data sets from the raw data provided. This requires students to use data analytics software (e.g., R, SPSS, MS Excel) to reduce and manage the dataset.
3) Prepare a report covering each of the questions.

The Submission Criteria

The submission will be maximum 2,000 words including all textual components of the document (only references are excluded from the word count). Do not use appendices, place any statistical output (properly formatted to include only relevant pieces of information) directly into the main body of the text.

Each question should be started on a NEW page. You should use Time New Roman size 12 font with 1.5 line spacing and 1 inch margins. Use headings to structure your answer to each question.

You must electronically submit your written assignment to Blackboard (Turnitin) so that it can be checked for plagiarism.

Group Size

The groups should consist of between 3 and 4 members. A penalty of 25marks will be subtracted from the final mark for the assignment if a submission has too few or too many group members. Students will be able to choose their own groups, and issues with regard to group work are not considered an appropriate basis for changing marks. If any problems in your group arise, or you have problems finding a group, please speak to me sooner rather than later to gain assistance – such issues are almost impossible to address in the week leading up to submission, so you should contact me early.

Marking Criteria

The criteria used to assess the group assignment will be as follows. This criterion will be applied to EACH QUESTION SEPARATELY with the total assignment mark then being the sum of the individual parts.

1) Choice of analytic approach – 20%

Having understood the question students should choose and JUSTIFY the analytic tool(s) they will use to obtain the insights needed to answer the question. The justification should be logical and should use appropriate references to support the argument.

2) Appropriate preparation of the data for analysis – 20%

Students will need to demonstrate that they can clean the data for the relevant variables for that question only prior to undertaking analysis. This cleansing will need to be done in knowledge of both the question being answered and the standards for obtaining high quality data. The issues you may need to consider include, removing data points irrelevant to the question at hand, addressing missing data, addressing erroneous data points, and addressing outliers.

3) Application of analytic tool – 20%

Students should undertake the analysis either by 'hand' or using 'data analytics software' (e.g., R, SPSS, MS Excel). Appropriate outputs should be reported in the body of the text (note: not all output is appropriate, only show those statistics useful to your analysis) and interpreted clearly and concisely.

4) Application of the insight – 30%

All questions asked in the assignment include a request for a 'comment on the result ...' or 'explain ...'. Addressing this application of the insight part of the question needs to be done clearly and fully. It should be addressed to a non-technical audience and USE THE RESULTS to inform the answer. Generic comments will not be marked highly, nor will recommendations that draw excessively on statistics that were not a part of the analysis; the comments should be very specific to the question asked. As you are given the task to explore business opportunities in

continental Europe, you need to draw insight from this dataset to inform your manager.

5) Structure and Presentation – 10%

Students should appropriately lay out the findings and discussion with headings and tables used to support their answers. Failure to follow the Submission Criteria (see above) will result in penalties here. The exact style and format is otherwise the responsibility of the group. There should be good use of the Harvard referencing system and general pagination across the document. Marks in this area can be achieved by delivery a logical and professionally presented document.