

Case Study 1: Are All Statistical Models Wrong? Understanding the Limitations and Usefulness of Statistical Models in the Popular Press

CLASSIFICATION OF ASSIGNMENT

- ❖ This is a Category B Assignment - A Group Assignment:
 - Your group may not receive help from anyone outside your group. All questions concerning this assignment should be addressed to your professor. It is an honor code offense to give help to other groups and individuals or receive assistance from other groups and individuals.
 - Groups are assigned by MSBA program.

INSTRUCTIONS

- ❖ Select a topic that has been discussed in the *popular press* online that can be solved by statistical analysis.
 - The term "*popular press*" refers to material written for the general public.
- ❖ The main goal of the assignment is to research a claim regarding the connection between two variables, and determine the limitations and usefulness of statistical analysis in understanding the problem. Without inserting your personal biases, critique the topic based on what you found and understand about the definition of statistical learning.
 - In doing so, from the articles you found, document the extent this topic has been analyzed using statistical models.
 - ❖ In the articles you found, was the dataset made public, and is easily accessible? Are the results clearly based on statistical models? Are the findings presented on this topic clear and concise?
 - Note: The authors of the articles you find may or may not have included p-values or the specific tests that were conducted, but still presented results or made claims anyway. The online articles you find may or may not follow the rules of scientific inquiry.
- ❖ In a team paper, include the following:
 - Your team name along with individual names listed of people that contributed.
 - A summary of the topic along with a clear description of the variables are being discussed.
 - Variables need to be clearly labeled as independent or dependent variables and defined according to the research found.
 - An argument for why this topic would classify as a ML example based on your understanding of the V's we discussed: Value, Volume, Variety, Velocity, Variability, and Veracity.
 - Present the usefulness of conducting statistical analysis to research this topic.

- Present any potential issues with the results presented in any of the articles found, including potential third variables, multicollinearity, the directionality problem, potential measurement error, and or selection bias.
- All citations including pdf copies of the articles used.
- ❖ One required zip file per group should be submitted via blackboard by the due date listed in the system. The zip should include the team paper and the articles found.

RUBRIC FOR PAPER (80%)

A (100) – This assignment is considered **exemplary**. A **timely** topic was presented that could be analyzed using statistical models. Arguments were **well-articulated** on the *Usefulness* and *Limitations* of statistics provided on the topic. **A clear and concise description** of the variables were included, and **ample research** was provided that summarizes the topic well. Paper Includes all necessary components including zip file of all research collected.

A- (90) – This assignment is considered **well-done**. A **timely** topic was presented that could be analyzed using statistical models. Arguments were **presented** on the *Usefulness* and *Limitations* of statistics provided on the topic. **A clear and concise description** of the variables were included, and **research** was provided that summarizes the topic well. Paper Includes all necessary components including zip file of all research collected.

B (82) - This assignment is considered **proficient**. A **timely** topic was presented that could be analyzed using statistical models. Arguments were **summarized briefly** on the *Usefulness* and *Limitations* of statistics provided on the topic. A **description** of the variables was included, and **research** was provided that summarizes the topic. Paper Includes all necessary components including zip file of all research collected.

C (75) - This assignment is considered **acceptable**. A topic was presented that could be analyzed using statistical models. Arguments were **summarized briefly** on the *Usefulness* and *Limitations* of statistics provided on the topic. **Detail was missing**. A **description** of the variables was included, and **research** was provided that summarizes the topic. **Not all assessments correct**. Paper Includes **most** of the necessary components.

F (60) – This assignment is considered **underwhelming**. A topic was presented that could be analyzed using statistical models, but **major detail was missing and or inaccurate**. Variables were not clearly defined, and not enough **research** was provided to summarize the topic. **Not all assessments correct**. Paper Includes **some** of the necessary components.

F (0) – No submission or topic unable to be analyzed using statistical models.

GROUP CONTRIBUTION AND ASSESSMENT (20%)

- ❖ Everyone in the group must contribute and write a portion of the team paper: introduction, summary and analysis of topic, conclusion, citations.
- ❖ Groups will be peer evaluated through teammates. **Participation and score through teammates will be incorporated into your final average.**