# **Yelp Business Data Analysis**

# The date you anticipate teaching the lesson during Fall semester 2020: October 22, 2020

# Number of hours required to teach to the lesson: 1 hour of class time

# Method of Lesson Delivery: online

# Teaching Assistants involved in delivering this lesson? (Y/N): N

# Brief lesson description: Students will think critically about how data collected from the Yelp app can be used in business and marketing.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Guiding Principles, Learning Objectives and Learning Outcomes** | **Discussion(s)** | **Assignment(s) and Assessment** | **Learning Resources + Media** | **Prerequisites and Expectations for Students** |
| **Learning Objectives**  After this lesson, students will be able to:   * Demonstrate an understanding of the properties, values, risks, and ethical considerations of data analytics in online business communities * Identify strategies to utilize data from online business communities. * Utilize data analytics techniques to create useful and relevant recommendations for organizations.   **Learning Outcomes**   * The students will describe the properties, values, risks, and ethical considerations of data analytics in online business communities explaining a sample of Yelp user data. * The students will frame useful questions that can be answered by specific data visualizations based on the data set and stated business problem. * The students will interpret visualization to inform business strategy. | How does a data dictionary help increase data understanding? What are the advantages to including a data dictionary with a dataset? Give an example of how data might be used incorrectly without a data dictionary.  What are some of the ways that similar businesses can be grouped, based on the columns in the data dictionary?  How can businesses take advantage of the data collected on Yelp users? Think about this question from a marketing perspective.  Is Yelp’s data collection a violation of privacy, or is it necessary to provide the best possible service to its users? See article d for further information.  Why might it be useful for students to hone their analysis skills? Which tools might you learn more about to handle large datasets like this one? For example, tools such as R, Python, and SQL are useful for handling large datasets.  Encourage students to share knowledge / experiences if they have used these tools or seen them used in your domain. | **Using an existing data set provided by Yelp, students will think critically about how data collected from the *Yelp* app can be used in business and marketing. Students will also make inferences from data visualizations and consider the data in context of privacy and ethics.**  **To prepare for this assignment:**   * **Review articles a. and d.**   **Assignment Part 1:**  **In-Class / Group Work Option:**  Work through the Yelp business analytics problems from the learning guide together in class. This guide incorporates polling, discussion questions, and the use of breakout rooms to engage students in understanding the material. One easy way to do this virtually is by screensharing the Power Point *Yelp\_presentation* and working through the slides as explained in the teachers’ notes.  Specifically, students will answer polling questions throughout the class and enter breakout rooms 3 times to answer open-ended questions.  Discuss different groups’ answers after students complete each section. When the lesson is complete, discuss the role of data science in business and marketing.  **Take-home Worksheet Assignment Option:**  Alternatively, have students complete the Yelp\_Worksheet\_student.  **Assignment Part 2 (optional):**  **Challenge Problem:**  Provide student groups with excel file of North Carolina Yelp data (NC\_yelp.xlsx). As a take-home activity, have them filter data according to instructions under “Challenge Problem” section in the Yelp\_Worksheet\_student worksheet. For added challenge, ask students to identify another category or city in NC and explain or create a visualization to communicate their findings.  **Assessment: TBD** | **Data:** https://www.yelp.com/dataset  Articles:   1. How Businesses Use Data: <https://hbr.org/2018/05/how-companies-can-use-the-data-they-collect-to-further-the-public-good> 2. Data Dictionaries and Metadata: <http://library.ucmerced.edu/data-dictionaries> 3. Data Dictionary Examples: <https://data.nal.usda.gov/data-dictionary-examples> 4. Ethics of AI: <https://github.com/generationai/Johns-AI-Intro-2020> | Knowledge of (including anything that might need to have been presented in a previous lesson): some understanding of bar charts and heatmaps may be helpful in this lesson.  Experience with (please be specific re requirements for computer science and math): familiarity with excel or a program with similar functionality will be necessary to complete the challenge problem.  Platforms required: Zoom or another online learning platform that allows for screen sharing, polling, and breakout rooms. Additionally, students must have access to excel or a program with similar functionality to complete the challenge problem. |