



Marvel Comics recently heard that Lobster Land has run successful winter carnivals at its location in Maine in previous years. This year, Marvel has approached Lobster Land regarding a partnership opportunity. Lobster Land has agreed to Marvel's terms. Jointly, the two entities will host a winter wonderland festival in Maine, at Lobster Land, from 21-24 December.



Your team will submit a written report via Blackboard. This will include:

- A written document that includes a link to your Tableau dashboard (or the .twb/.twbx file itself) and your write-up/description of the dashboard.
- A PDF with an accompanying .ipynb that includes your results for the following parts:
  - Summary Stats
  - Segmentation and Targeting
  - Conjoint Analysis
  - Forecasting
  - Classification
  - A/B Testing
- Your case analysis -- this should be written in a document of between 2-3 pages
- Your conclusion -- this can either come in a Markdown cell at the end of your Jupyter Notebook or it can be its own document that you upload into the .ZIP.

During the last week of the semester, your team will deliver a 15-minute presentation in class to your Professor and TA.

Some things to keep in mind:

- As you move through the various tasks, remember to “call ‘em like you see ‘em.” If you see results that aren’t pretty, it does not necessarily mean that you did something wrong, or that there is a problem with the dataset or the software.
- Some tasks in this project are very similar to things that we have done in homework assignments, whereas others are unique. Take advantage of *all* your available resources...but keep in mind that the teaching team will remain at “arm’s-length” distance from all project teams.
- Assumptions are okay -- if you make an assumption anywhere along the way, you can just state it so that your Professor can see why you took some particular step.
- The sections are weighted differently – keep this in mind as you go.
- A rubric is available on Blackboard.

**Data Visualization (10 points):** Marvel is perhaps best-known for its very popular series of movies. However, Marvel originally gained fame for the comic books that it has published. Marvel has gathered data about those comic books, and now they’re coming to Lobster Land to ask for some help with analysis.



Using Tableau, build a dashboard that includes anywhere from 4 to 6 visualizations created from the variables in the *marvel\_comics* dataset.

Include a 2-3 paragraph description of your dashboard that talks about the plots you made and some of the valuable takeaways and insights that they may provide for Marvel management. In your analysis, you may also mention any limitations associated with the dataset or its interpretability.

**Summary Stats (5 points):** In Python, conduct some exploratory data analysis of the *marvel\_movies* dataset.

You may want to present some summary statistics of the entire dataset, but should also consider some groupings of variables, using `groupby()` or `pivot_table()` from pandas. Create anywhere from 4 to 6 total summary statistics for this section.

In a paragraph, describe your findings. What did you learn about the dataset? Mention any insights that might be particularly useful or valuable for Marvel management. Use a markdown cell to write this paragraph. In your analysis, you may also mention any limitations associated with the dataset or its interpretability.

**Segmentation and Targeting (25 points):** Lobster Land is planning to heavily emphasize local marketing for this event. Because travel in this region can be difficult during winter months, Lobster Land plans to focus most of its outreach efforts on households near Portland, the largest city in Maine.

Here's where your team can be a huge help to Lobster Land: The park has recently obtained a dataset with information about a large number of area households, but the park doesn't know how to begin analyzing it. The dataset is named *regional\_families.csv*. Park management has heard about something called clustering, but they're not quite sure what it is, or how it works....and they think your team can help out.

You may wish to use either k-means or hierarchical clustering for this task. To perform the actual clustering, use only your numeric variables (but when you analyze your clusters, you can include observations about the categorical factors).

Once you have built your clustering model, use anywhere from 4-6 visualizations that help to communicate information about your model. The visualizations should depict information about your clusters that you can clearly explain, and that park management can understand. So stay away from things like PCA and t-SNE!

Name each one of your clusters, and include a few sentences describing/explaining the name that you chose for each cluster.

Finally, for each of your clusters, talk about targeting. In a couple of sentences per group, how should park management reach each of these segments? In your answer, you should think about prioritization, too – into which clusters should Lobster Land invest its most significant efforts?

Use markdown cells for the cluster names and the targeting section. Also, use a markdown cell to describe the process that you used for arriving at the number of clusters for your model.

As you analyze the dataset, bear in mind that this data is based on a *sample* of households from across the region. Do not expect the variable values to be reflective of the overall population.



### **Conjoint Analysis & Memo Section (20 points):**

Winters in Maine tend to be quite cold, but Lobster Land has a potential concept in mind that might help to keep visitors warm: Barbeque on the Lake.

The food vendor will need to travel from western Massachusetts to Portland, ME, so the vendor is requiring Lobster Land to select just one Starter, two Main Dishes (one each from two groups of choices), one side, and one dessert. Lobster Land management is feeling overwhelmed by the available options.

All options come with unlimited soda and dipping sauce -- that's why drinks and dipping sauce do not appear on the menu. Lobster Land also hopes that local beer vendors can sell beverages that will go well with the food choices shown below.

For Starters, the options are: Fried Chicken Tenders, Crabcakes and Shrimp, Sticky Chicken Tenders, and Jumbo Shrimp Cocktail.

For Main Entree, the choices are: BBQ Ribs, Grilled Sausages, Smoked Turkey Legs: Grilled Salmon, and Grilled Portobello Mushrooms.

For Salad/Soup, the choices are: Winter Salad, Potato Salad, Chili, and Beef Stew.

For Side, the options are: Mac and Cheese, Mashed Potato, and French Fry Basket

The Dessert Options are: S'mores, and Apple Pie a la Mode

The *bbq\_lake.csv* file includes consumer ratings for each proposed bundle. The avg\_rating variable gives the mean of all ratings for each particular bundle, as determined by more than 2000 total survey responses per bundle. When consumers gave these ratings, however, they did so in a cost-neutral way -- they were asked to simply rate the bundles based on taste/preferences.

The *vendor\_costs.csv* dataset contains information about the per-serving cost associated with each option that has been presented to Lobster Land.

Lobster Land management has decided that it will charge a flat \$15 fee for all visitors who enter the Barbeque Tent at the Winter Wonderland.

Using the dataset *bbq\_lake.csv*, perform a ratings-based conjoint analysis of the bundle options. Write a one-page executive summary to Lobster Land. In the memo, be sure to clearly state your recommendation for the hot pot options that Lobster Land should go with.



**Forecasting Total Revenue (3.75 points):**

Through an online search, you can find recent annual reports published by Disney, the company that owns Marvel Comics. The company trades publicly, with the ticker symbol DIS.

Using several years' worth of data, and any forecasting method that your team deems appropriate for this purpose, generate a forecast for DIS's 2023 end of year total (gross) revenue. Use any forecasting tool available in Python to accomplish this.

In a markdown cell, write one paragraph that describes your process and results. In addition to the DIS revenue data from previous years, you may also wish to consider any other factors -- but keep the weighting of this section in mind here.

In a markdown cell, write one paragraph that describes your process and results.

**Classification (15 points):** In this section, you and your teammates will use the *carnival\_visitors* dataset to help Lobster Land better predict whether a particular household's primary goal for attending a winter carnival will be 'indoor' or 'outdoor'.

You can assume that those who prefer 'indoor' will be primarily interested in things like holiday-themed shows, craft fairs (inside heated tents!), and indoor games or eating. Those who prefer 'outdoor' will be primarily interested in outdoor concert performances, physical activities (like ice skating or snowmobile races), or hands-on activities like Do-It-Yourself ice sculpting.

You may use any classification method that you have seen anywhere in AD654 material. Show the results of your model, and the steps that you used to build it.

Write 3-4 thoughtful paragraphs about the conclusions that Lobster Land management can draw from this model's results, and/or marketing approaches that they can apply based on the outcomes. In this assessment portion, you may wish to point out any specific details about your model to help park managers better understand it.



**Strategic Memo (10 points):** You and your teammates should each read the case, “[Marvel Enterprises, Incorporated](#)” *This case will not be posted to Blackboard.*

Lobster Land is thinking about expanding its brand. Lobster Land is inspired by the example that Marvel has set, and hopes to learn from Marvel’s example.

After reading the Marvel Enterprises case, write a memo of 2-3 pages in which you address *some* of the following questions (you do not need to cover all of them):

- *How can we effectively leverage popular characters or intellectual property (IP) in our theme park?*
  - *What new attractions or shows can we develop that are themed around these characters or stories?*
- *In what ways can we expand our brand beyond the theme park to engage customers year-round?*
- *What potential partnerships (e.g., with movie studios, TV networks, or toy manufacturers) could enhance our brand and attract more visitors?*
- *What unique merchandise can we create that ties into our theme park's unique attractions or characters? How can we optimize the placement and marketing of merchandise within the park to maximize sales?*
- *What special events or seasonal programs can we introduce to attract visitors during off-peak times?*
- *What feedback mechanisms can we implement to gather insights from our visitors?*
- *Apart from ticket sales, what other revenue streams (like memberships, VIP experiences, or educational programs) can we develop?*
- *How can we balance revenue diversification with maintaining a quality visitor experience?*

For this section, consider this case to be a “closed universe.” You may bring your opinions, and your teammates’ opinions, into the case, but you will not need to do any outside research, and should not cite any outside research here. You can use some creative freedom in the way you describe Lobster Land and its current operations. Of course, Lobster Land is a fictional place, so you are free to fill in any gaps about Lobster Land with reasonable assumptions.

Very good answers for this section will include specific recommendations and ideas for Lobster Land, based on findings/lessons from the Marvel case. Avoid lengthy summarizing of details from the case, but do mention relevant details from the case as you formulate your guidance for Lobster Land.

You and your team may think of this case as its own “island” -- the questions here are distinct from any other sections of the project.

For this section in particular, be careful with chatGPT *abuse*. While you may use any source (including large language models) for inspiration, a copy/paste job from chatGPT is \*not\* the way to go with this section. Be creative, be genuine, be specific, and make the call.

**A/B Testing (10 points):** Lobster Land is planning to include some high-quality ice sculptures at this year’s Winter Wonderland event.

Lobster Land temporarily purchased the digital rights to three photographs, shown below. To test out the popularity of each photograph, they sent e-mails to Lobster Land fans, and randomly chose one of these three images to include in the e-mails. From top to bottom, the first picture shown here is known as “Iron Man.” The second picture shown is “Spider Man”, and the third picture shown is “Captain America.”







Lobster Land has obtained some data about the photos, and now they've come to your team for help. Lobster Land must choose just *one* of these three pictures for its 2023 Winter Wonderland marketing campaign.

Use appropriate statistical test(s) for this situation, and deliver recommendations for the park regarding the pictures.

In a markdown cell, write one paragraph that describes your process and your results for analyzing the results of these A/B Tests.\* Which picture should Lobster Land use?

\* Even though there are several differences between these pictures, we can still consider this comparison an A/B Test. The results will let us understand the differences from picture to picture, even if they won't enable us to draw more detailed conclusions.

**Conclusions (1.25 points):** Here, write 1-2 paragraphs to wrap everything up. Do not merely describe your findings as if you were summarizing a lab report. Instead, emphasize your team's insights about how these findings could be helpful/meaningful for Lobster Land, or reflect on your team's process. This could also be a chance to talk about other things for Lobster Land to consider, besides

what has been mentioned here. For this section, you may simply use a Markdown cell, or you may write in a separate document that you include in your ZIP -- either way is fine.