

NAME: Owen Sullivan

OBJECTIVES: The purpose of this project is to introduce the statistical software JMP. Upon successful completion of this project, you will be able to...

- Identify the type of variable in the context of a problem
- Create graphical displays, frequency tables, and relative frequency tables in JMP for qualitative data
- Make and support decisions on the best way to display data
- Use the output provided by JMP to answer relevant questions.

Delectable Delights is a large consumer food manufacturer selling its products in retail stores nationwide. You have landed your first job after graduation from Clemson in their advertising division. Since you took statistics as a part of your coursework, you are often called upon to perform data analysis for the advertising division, as well as other divisions of the company.

DIRECTIONS: Answer the following questions using complete sentences as though you were presenting your analysis to the employees of Delectable Delights. Please provide any appropriate output and/or screenshots from JMP. Instructions for creating several types of graphs or tables and statistics can be found on Canvas in the file **JMP Instructions.docx**. Paste your answers and any output into this document.

1. (65 points) Delectable Delights would like to feature some of its products in major motion pictures. To assist in selecting possible movie projects to approach, your boss, Ray Holtz, would like you to review movie genres to see which is most popular. Since Delectable Delights uses the software JMP, you decide to look into the Hollywood Movies data set that comes with JMP. You can find the file **Hollywood Movies.jmp** on Canvas.
 - (a) Observe the column named “Genre”, then classify the variable “Genre” as: 1. qualitative or quantitative, 2. discrete, or continuous, or neither, 3. nominal, or ordinal, or interval, or ratio. (10 pts)
 1. **Qualitative**
 2. **Neither**
 3. **Nominal**
 - (b) Using only the Genre column in the Hollywood Movies data set, create the following graphical displays in JMP. Copy and paste your tables and graphs into the space provided.
 - i. **Frequency / Relative Frequency Table** (10 pts)
Step 1, create a frequency / relative frequency table.

Follow the steps on page 2, which are copied below, of the JMP Instructions under **Frequency table, relative frequency table, and histogram** to create a frequency/relative frequency table for the variable “Genre”.

Frequency table, relative frequency table, and histogram:

1. Look at the toolbar on the top and find Analyze. Click on it.
2. Click on Distribution, the first option.
3. Highlight the column of interest (In this problem it will be Genre) -> Click Y, Columns -> Click OK. Alternatively, you can drag the column to Y, Columns.
4. Click on the red arrow next to “Distributions” and select “Stack” to switch the histogram to be horizontal.

Step 2, modify titles: Change the column titles of the table to Genre, Frequency, and Relative Frequency instead of Level, Count, and Prob. To change a title, simply double-click the title and enter your own.

Step 3, copy/paste your table: Right click on the table, and select **Copy Table** from the drop-down menu, then paste it here.

Don’t close the JMP window. We will use it for bar chart next.

Genre	Frequency	Relative Frequency
Action	32	0.23529
Adventure	1	0.00735
Animation	12	0.08824
Comedy	27	0.19853
Drama	21	0.15441
Fantasy	2	0.01471
Horror	17	0.12500
Romance	11	0.08088
Thriller	13	0.09559
Total	136	1.00000

ii. **Bar Chart** (10 pts)

Step 1, create a bar chart: We almost have a bar chart in the same JMP window when we created the frequency/relative frequency table. By default, JMP does not separate the bars. Make sure the bar chart is horizontal, otherwise check the JMP Instructions above.

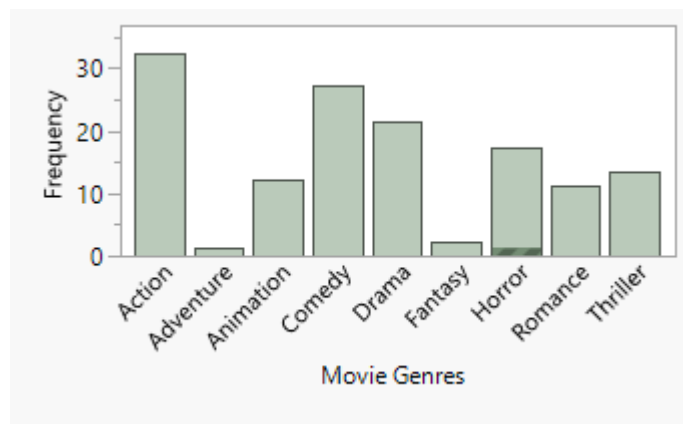
Step 2, modify axis/labels:

For y-axis, click the red triangle to the left of the column heading (in this case it is Genre). Select **Histogram Options >> Count Axis** or **Prob Axis**. This will place either the frequency or relative frequency axis on the bar chart depends on what you want to plot. Once more click the red triangle and select **Display Options >> Axes on Left** to change the location of y-axis. Change the label of the y-axis to Frequency instead of Count, or Relative Frequency instead of Probability by single clicking on the default label.

For x-axis, right click on the movie genres on x-axis and select **Add Axis Label**, then enter the name “Movie Genres”.

Step 3, aesthetics: Click the red triangle again left to “Genre” and select **Histogram Options >> Separate Bars**. If you would like to change the color of the bars, right click in the white space around the bars and select **Histogram Color**.

Step 4, copy/paste your chart: Right click on the white space around the bars and select **Edit >> Copy Graph**, then paste it here. (There are instructions on page 6 of the JMP Instructions on how to use the Graph Builder feature to make a bar graph. JMP often has more than one way to create pictures of your data.)



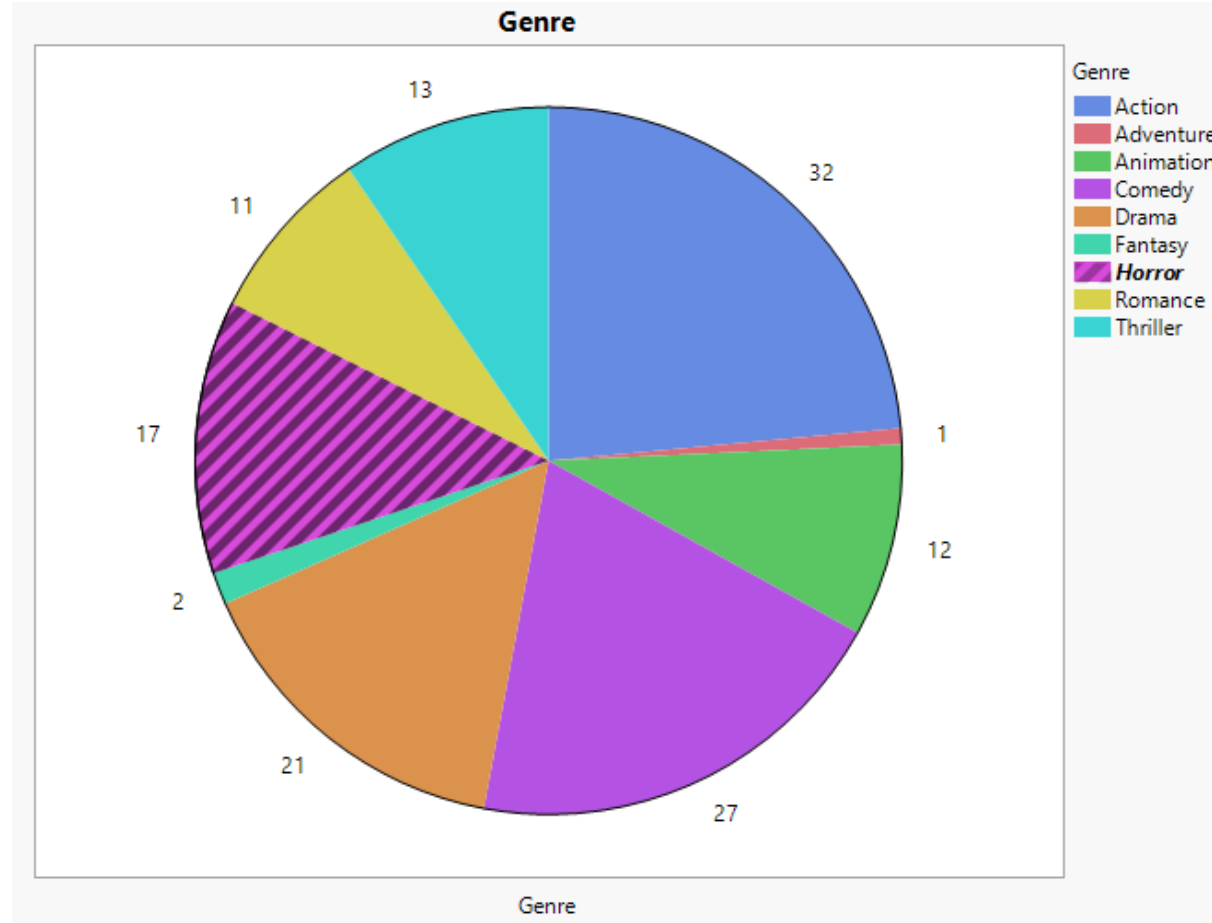
iii. Pie Chart (10 pts)

Step 1, create a pie chart: Follow the steps on page 5 of the JMP Instructions to make your pie chart. The instructions are copied below.

1. Click on Graph in the toolbar on top.
2. Click on Graph Builder, the first option.
3. Click **Pie Chart** picture.
4. Drag the variable of interest into white zone in the middle.

Step 2, add counts or percentages: Use the **Label** dropdown on the bottom left of the Graph Builder window. Choose “Label by Value” or “Label by Percent of Total Values” depending on what you want.

Step 3, copy/paste your chart: Copy and paste your chart here. To copy the pie chart, right click on the chart, go to **Edit >> Copy Graph**. This should both copy the graph and the legend.



(c) You need to make a short report to Ray on the Genre that is most popular.

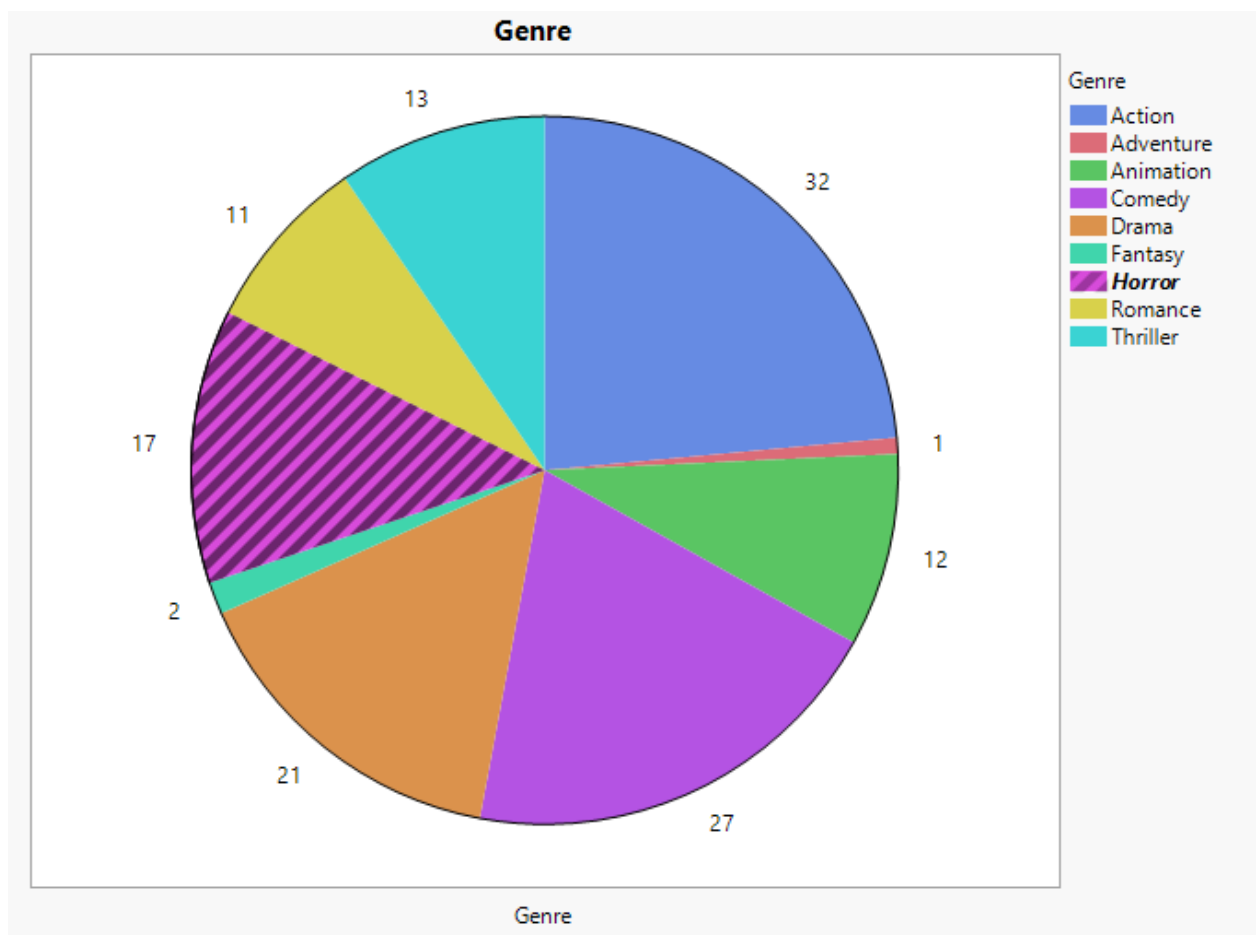
- i. **Which** of the tables/graphs that you made in Part (b) would you like to use? **Why** did you select that table/graph? (There is no one right answer to this question. You need to select the type of graph that you believe quickly conveys the count or proportion or percentage of each genre). (5 pts)

I would use the Pie Chart for my report, because it shows the distribution of movie genres in a way that is easier for someone to understand without looking at the numbers or understanding the background context (it uses colors and proportionately-sized areas of the circle).

- ii. Calculate the proportion of the most popular movie genre **by hand**. Since the data given is a sample of the movies produced in a particular year, this will be a calculation of **sample proportion**. Include the **symbol** for sample proportion, show your work by **filling in the formula** for sample proportion, then get the **result**. (10 pts)

Action movie genre: $\hat{p} = x/n = 32/136 = 0.23529412$

- (d) Using the graph or table you chose in Part (c), **report** to Ray the most popular movie genre. **Explain** your answer based on the graph or table you selected. (10 pts)



According to the Pie Chart above, the largest proportion amongst movie genres is Action. The proportion of people in the sample of 136 that prefer Action movies is 0.235. This means that, of the 136 sampled, 32 preferred Action movies.

2. (35 points) Let's separate the movie Genres based on whether they are produced by an independent studio or one of the major studios. We will see if your answer in the previous question about which Genre is most popular depends on the movie studio.
- (a) Use JMP to determine how many movies are produced by an independent studio and how many are produced by a major studio. **Paste** the result generated by JMP here, and **briefly introduce** how you used JMP to find the result. (10 points)

Independent Studio	Genre								
	Action	Adventure	Animation	Comedy	Drama	Fantasy	Horror	Romance	Thriller
NO	28	1	12	22	14	1	12	7	7
YES	4	0	0	5	7	1	5	4	6

Using JMP, I tabulated the data set by comparing Genre against the Independent Studio variable. I dragged the Genre variable to the top row, and Independent Studio to the first column on the left.

- (b) Use the **Graph Builder** feature in JMP to make a bar chart for Genres that separates the movies by studio. (10 points)

Step 1, create a bar chart: Go to **Graph >> Graph Builder**. Select the **bar graph tab** at top then place the **Genres** column on the horizontal axis.

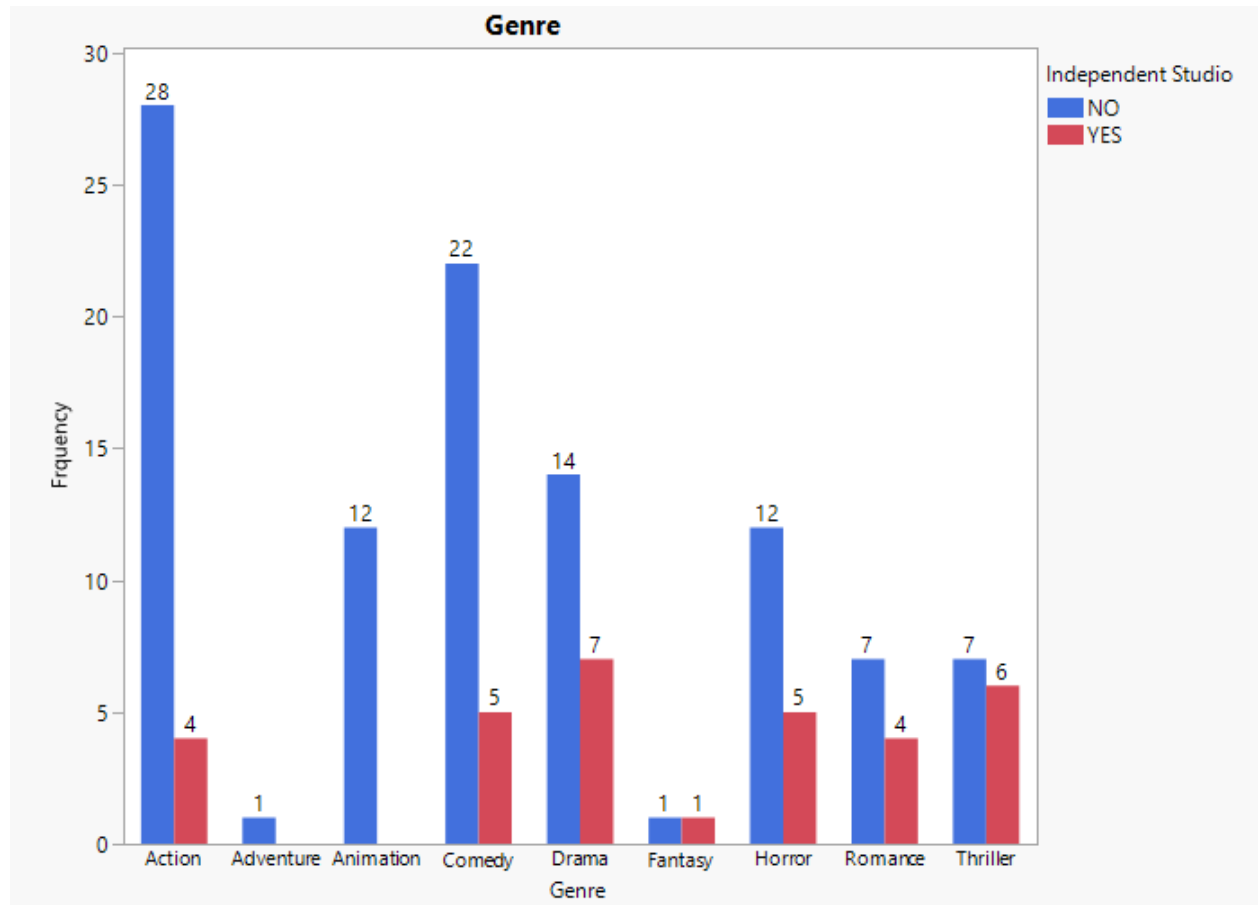
Step 2, separate bars by groups: To separate the Genres by studio, place the **Independent Studio** column into the box on the right of the graph labeled **Overlay**. The default graph will be a side-by-side bar chart of the movie genres with the frequency on the vertical axis.

(Optional) Step 3, explore other settings:

Your goal is to make a graph that quickly conveys the most popular genre for each type of studio. Experiment with the different types of bar charts JMP produces. On the lower left of the Graph Builder window are various options for the graph. Change the **Bar Style** options and choose the one you like the best. Also experiment with the **Summary Statistic** options. The default statistic is **Mean**, which provides you with the frequencies on the vertical axis. Experiment and see what happens when you change the **Summary Statistic** from **Mean** to **% of Total**, or some of the other options. Be careful that you understand how the information is being displayed. You may need your answer from Part (a) to help you.

Step 4, modify labels/axis: Right click on the y-axis and select **Add Axis Label**, then enter the name "Frequency".

Step 5, copy/paste your chart: Right click on the chart, go to **Edit >> Copy Graph**, paste it here.



- (c) Suppose that Delectable Delights wants to advertise with major studios. Which is the most popular genre? Write an answer to Ray and **support** your answer using the Bar Graph you selected for Part (b). (5 points)

The bar graph above shows that Action movies are the most popular genre of movies produced by major studios, indicated by the fact that the largest value in blue (which represents the "NO" value for the Independent Studio variable) is Action.

- (d) Suppose that Delectable Delights wants to advertise with independent studios. Which is the most popular genre? Write an answer to Ray and **support** your answer using the Bar Graph you selected for Part (b). (5 points)

The bar graph above shows that Drama movies are the most popular genre of movies produced by independent studios, indicated by the fact that the largest value in red (which represents the “YES” value for the Independent Studio variable) is Drama.

- (e) If you notice that the most popular genre from independent studios (the answer from (d)) is different from that of major studios (the answer from (c)), try to explain the reason. (The answer may vary. You will get credit as long as it's reasonable.) (5 points)

The most popular genre from major studios is Action, but the most popular genre from independent studios is Drama. This could be due to the fact that major studios have significantly more funding for Action movies with thrilling CGI, but independent studios instead allocate their budget for better writing and thus produce more captivating Drama movies.