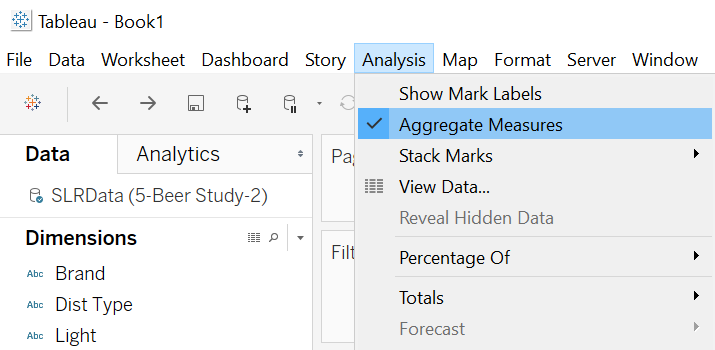
# Advanced Data Visualization – Lab #4

|  |  |
| --- | --- |
| Provided file(s): | * Lab04.docx * SLRData (5-Beer Study-2).csv |
| Submission file(s): | * Lab04.docx * Lab04\_1.twb * Lab04\_2.twb |

Work in your groups. Only one member of your group should submit. Only the last submission submitted before the deadline will be marked.

## **Part I: Scatterplot**

1. Open Tableau Desktop.
   1. Choose Open > File and select “SLRData (5-Beer Study-2).csv”.
   2. Drag “Calories” to columns and “Pct Alcohol” to rows shelves. You should get just one point.
   3. From the Analysis menu, deselect “Aggregate Measures” (as in the image). You should see a scatterplot now.

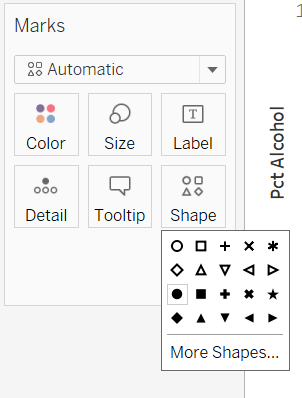


* 1. Change the title to

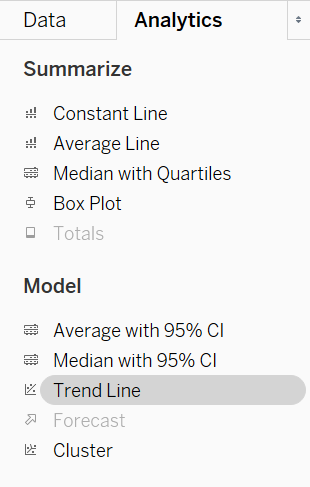
**Group #:**

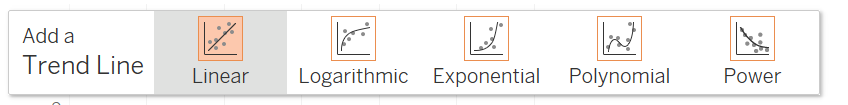
**Do Calories Affect Alcohol Content in Beer?**

This simple scatterplot shows the linear relationship between calories and percent of alcohol.



* 1. Use the Shape card (from Marks card) to change the shape of data points to filled circles.

* 1. From the Analytics tab, drag “Trend Line” onto the worksheet. You should see the different trend lines. Drop onto “Linear”.



The resulting trend line is a best fit linear model, using the method of simple linear regression.

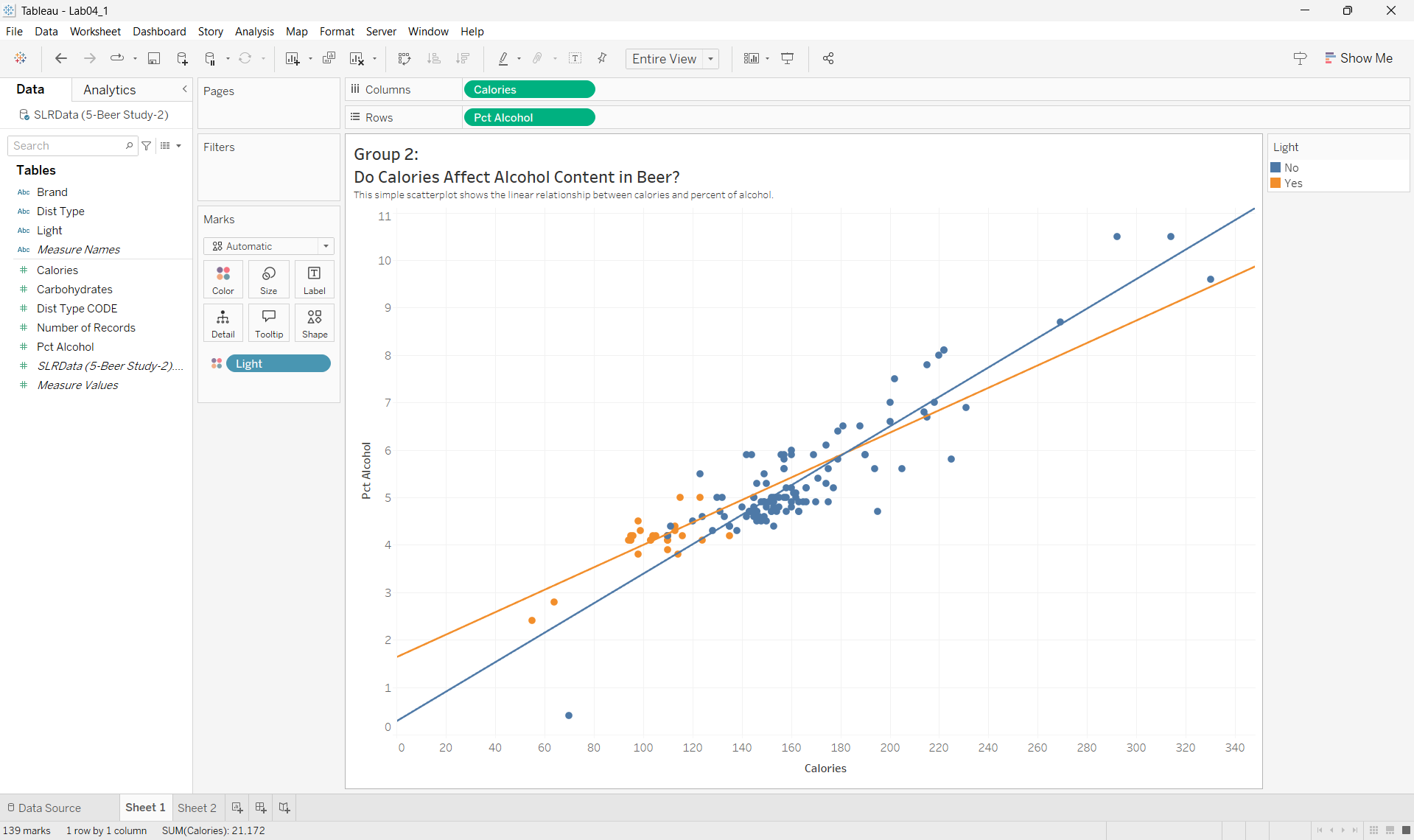
* 1. Right click the trend line and choose “Describe Trend Line”. Copy the information displayed and then paste it below:

|  |
| --- |
| P-value: < 0.0001  Equation: Pct Alcohol = 0.0275028\*Calories + 0.952598      Coefficients  Term Value StdErr t-value p-value  Calories 0.0275028 0.0011233 24.4849 < 0.0001  intercept 0.952598 0.177656 5.36203 < 0.0001 |

* 1. Right click the trend line and choose “Describe Trend Model”. Copy the information displayed and then paste it below:

|  |
| --- |
| Trend Lines Model    A linear trend model is computed for Pct Alcohol given Calories. The model may be significant at p <= 0.05.    Model formula: ( Calories + intercept )  Number of modeled observations: 139  Number of filtered observations: 0  Model degrees of freedom: 2  Residual degrees of freedom (DF): 137  SSE (sum squared error): 43.6047  MSE (mean squared error): 0.318282  R-Squared: 0.813988  Standard error: 0.564165  p-value (significance): < 0.0001    Individual trend lines:  Panes Line Coefficients  Row Column p-value DF Term Value StdErr t-value p-value  Pct Alcohol Calories < 0.0001 137 Calories 0.0275028 0.0011233 24.4849 < 0.0001  intercept 0.952598 0.177656 5.36203 < 0.0001 |

* 1. From the Data tab, drag and drop “Light” from Dimensions onto the Color card. You should now get two trend lines of different colours, for the different types of beer. Paste a screenshot of your graph below:



## **Part II: Matrix of Scatterplots**

1. Add a new worksheet to the file you were working on in Part I.
2. Drag “Calories”, “Pct Alcohol” and “Carbohydrates” onto both rows and columns and deselect “Aggregate Measures” (as you did in section (c) of Part I).
3. Drag the “Light” dimension onto the Color card.
4. Give the graphic a descriptive title, then paste a screenshot of your visualization below:



1. Save your tableau file as **Lab04\_1.twb**.

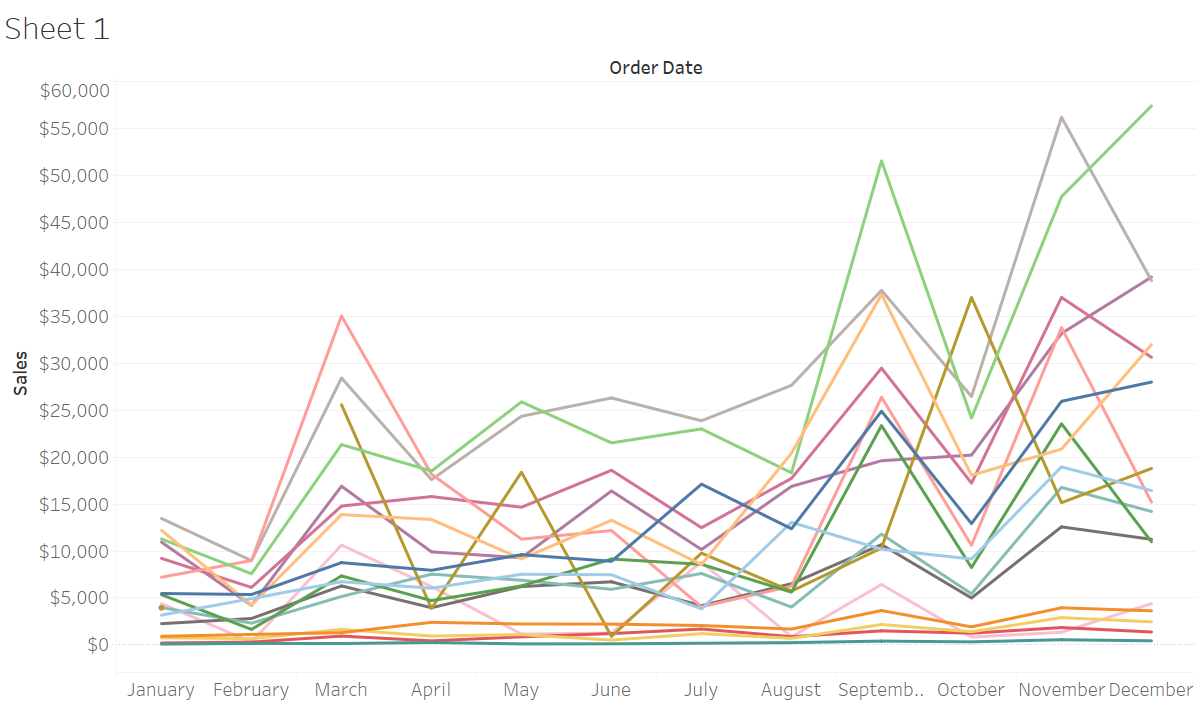
## **Part III: Analysis**

Write a detailed analysis of the Beer data below, based on your visualizations:

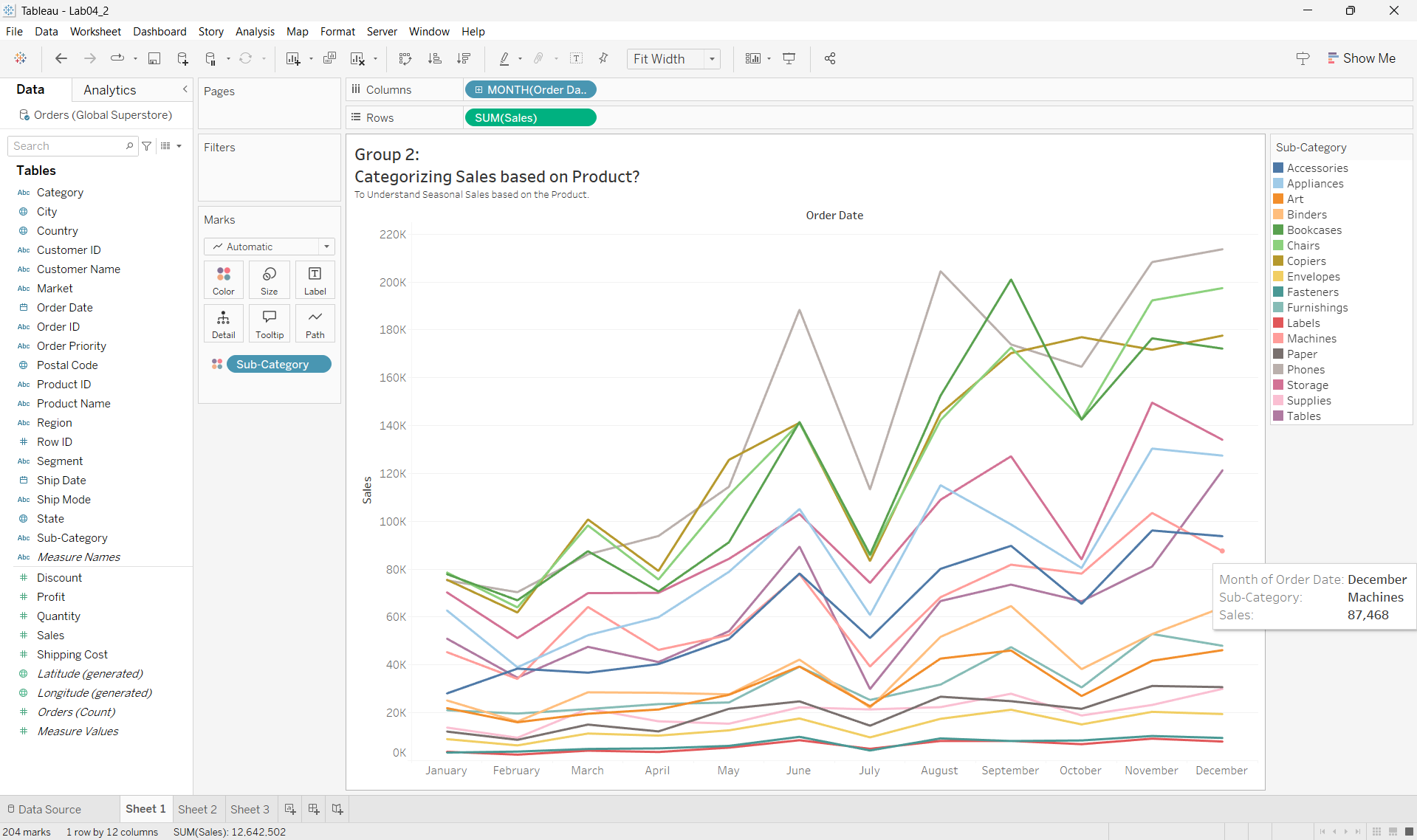
*In the initial visualization, we can determine in most cases the more alcohol, the more calories. In the second visualization, we can determine that the percent of alcohol and the amount of carbohydrates directly impact the amount of calories in beer. While the percent of alcohol and the amount of carbohydrates have a weaker relationship.*

## **Part IV: Heat Maps**

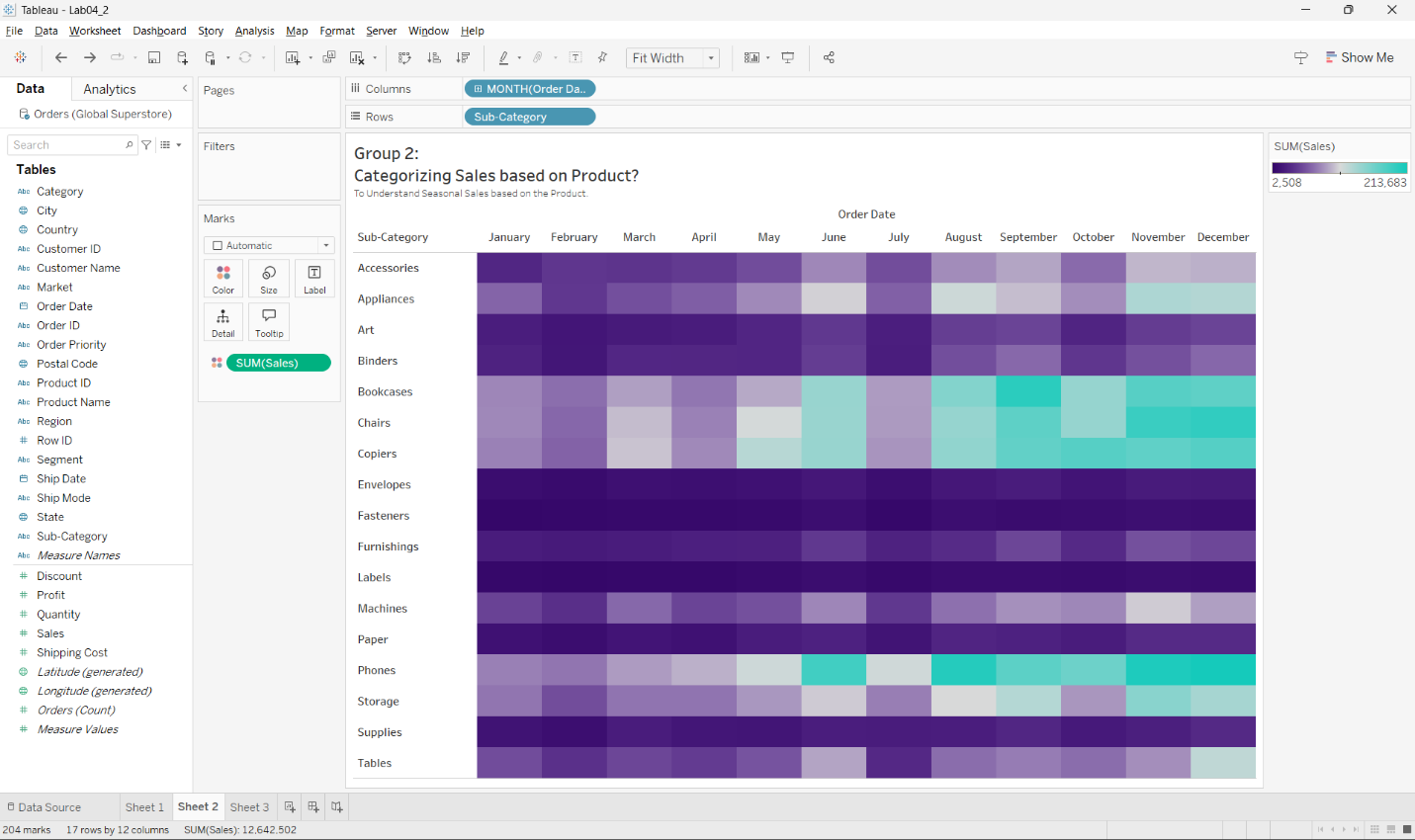
1. Open Tableau Desktop.
   1. Open the Sample superstore dataset.
   2. Drag “Order Date” dimension onto the Columns card. Right click and choose ‘Month’.
   3. Drag “Sales” onto into Rows.
   4. Drag the Product “Sub-Categories” dimension onto the Color cards. You should get this plot: (Ctrl+Shift+B if you want to make it wider)



1. As you can see, this is a much-cluttered plot! Open a new worksheet to create a heat map to present the same information.

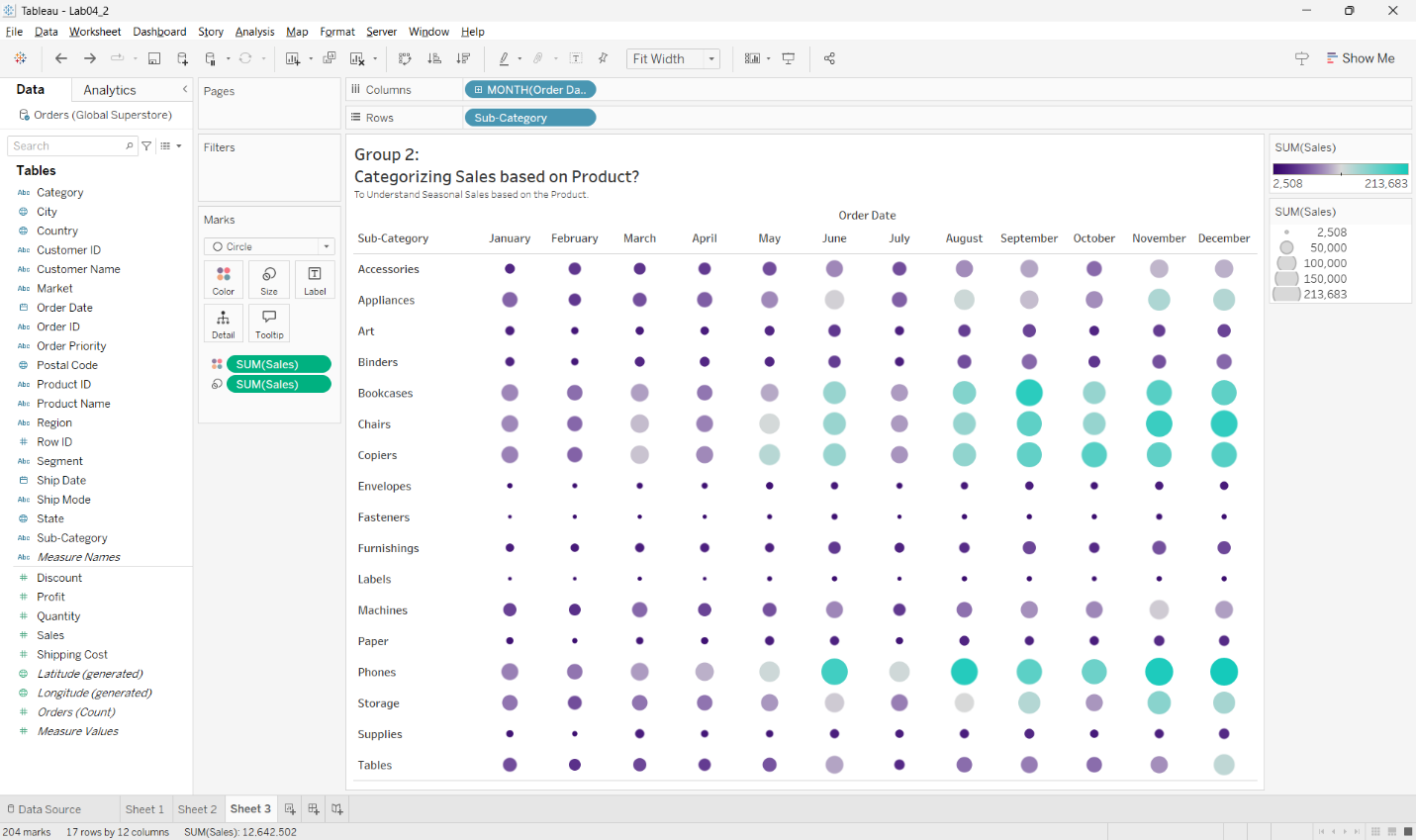


1. Drag “Order Date” dimension onto the Columns card. Right click and choose ‘Month’. Then right-click and choose ‘Discrete’ (it might already be selected).
2. Drag Product “Sub-Categories” dimension onto Rows.
3. Now drag “Sales” onto the Colors card. You should see a heat map now.
4. From the Color card, choose ‘Edit Colors’ and choose a different color palette. Note the diverging vs. sequential palettes. Experiment with both choices, then describe the difference below:
5. Which type of color palette (diverging or sequential) is most suitable in this case (for sales)? What would be an example of an application where the other choice would be more appropriate?
6. Finalize the color palette with one that you think works well.
7. Edit the title to something appropriate, and paste a screenshot of your heat map below:



## **Part V: More heat maps**

1. Add a new worksheet to the file you were working on in part IV.
   1. Repeat steps (f) and (g) from part IV.
   2. From Marks card, choose Circle.
   3. Drag and drop Sales onto the Color card.
   4. Now drag and drop Sales onto the Size card.
   5. Edit the title and paste a screenshot of your heat map below:



* 1. Save your tableau file as **Lab04\_2.twb** and submit.

## **Part VI: Analysis**

Write a detailed analysis of the Superstore data below, based on your visualizations:

*When we Look at the Heat-maps we can Understand per Category Sales in each Month, Doing so we can Understand if the Item’s Sold is Seasonal or Not, Which inturn can help the Shop to Optimise their Inventory & Hence Avoid Losses.*