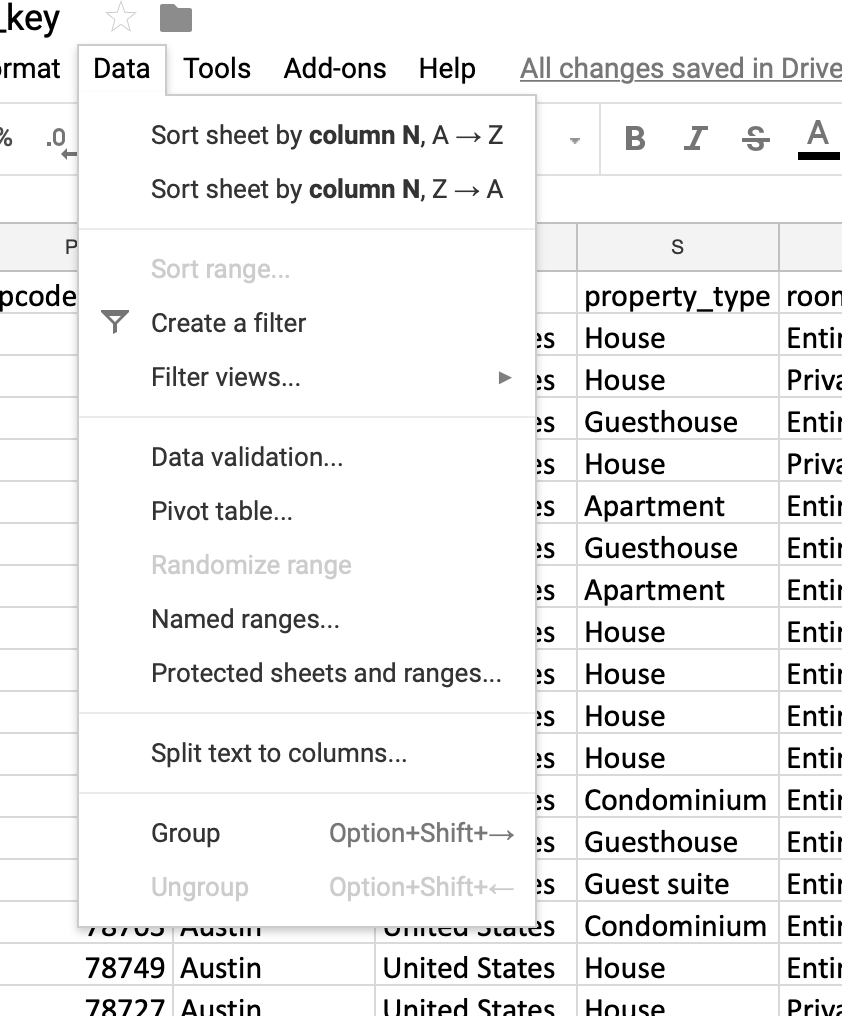
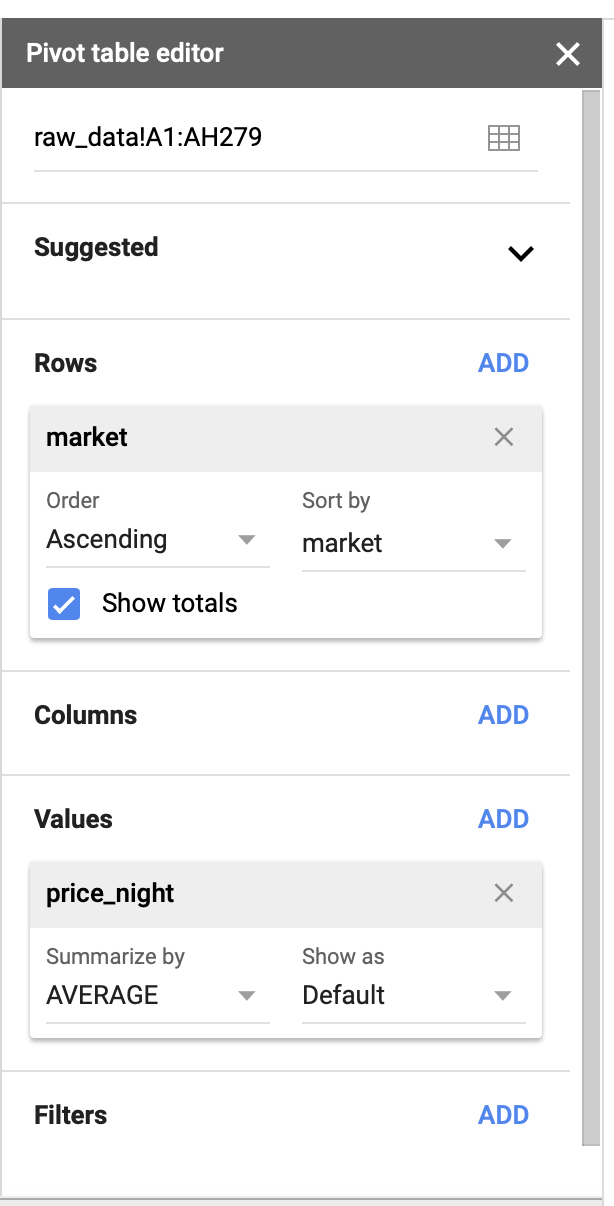
**Descriptive Statistics Questions**

1. Average listing price per city?

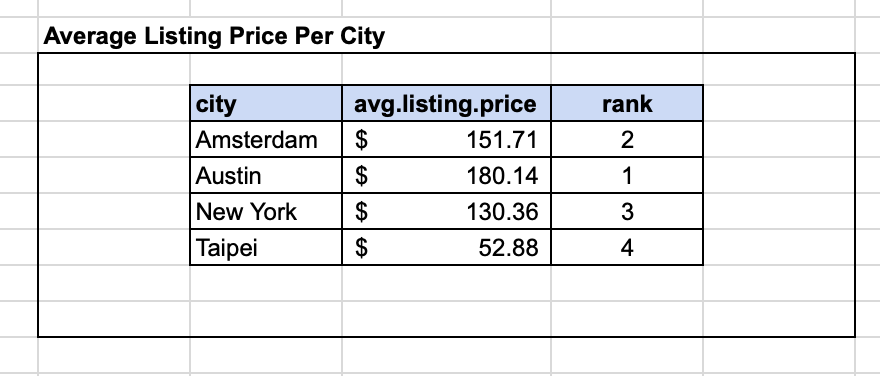
Select the raw\_data tab and insert a Pivot Table from the Data Tab in Google Sheets.



The Pivot Table will now populate a new sheet in which you can add the variables from the raw\_data. In the Pivot Table Editor, choose market for the Rows Shelf & Average price\_night for the Values Shelf.



Copy and paste the data you have calculated into Q1 of the worksheet under avg.listing.price



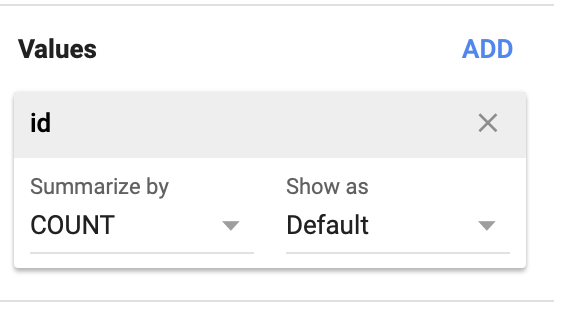
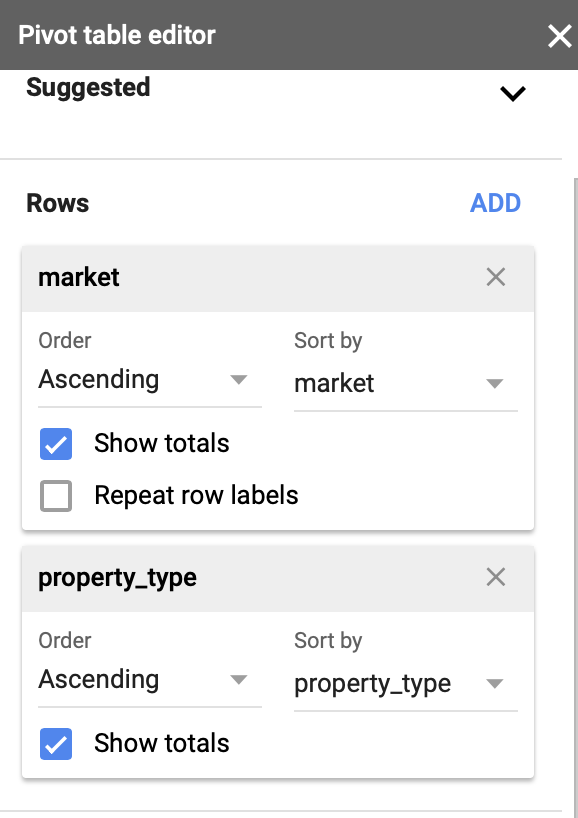
**Descriptive Statistics Questions continued**

1. What kind of property types are available in each city?
2. Build out a visualization for property types by city.
3. What insights can you draw from this data to inform the analytics team?

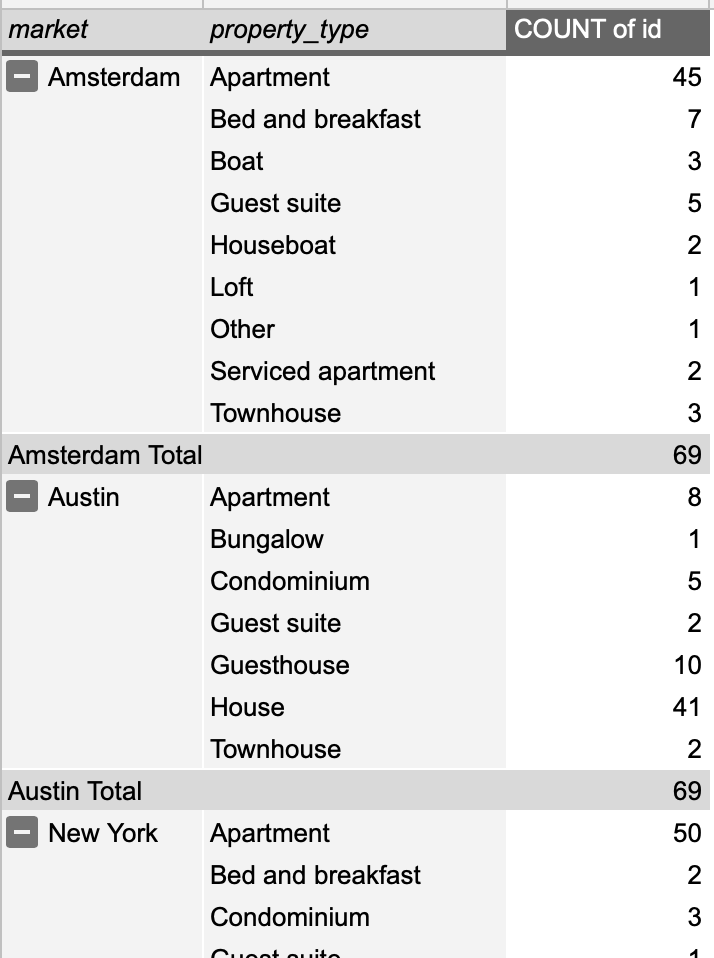
Congrats! You have completed your first solution as a business analyst. Now, let’s keep going!

Go back to the Pivot Table Editor and remove price\_night from the Values Shelf.

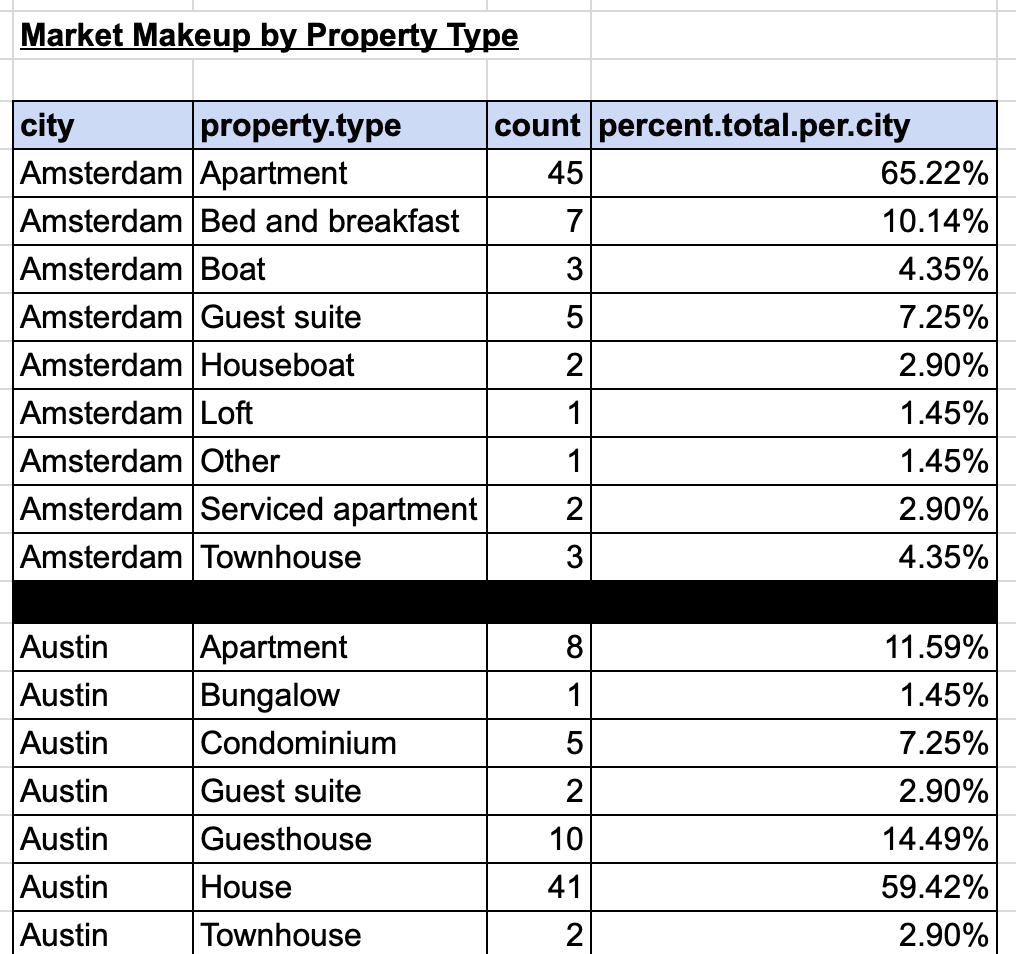
Add property\_type to the Rows Shelf and COUNT the number of id variables in the Values Shelf as shown below.



After you have added the appropriate variables to your Pivot Table Editor, your information in your table should look like this. This table represents the total number of listings in each market broken down by property\_type.



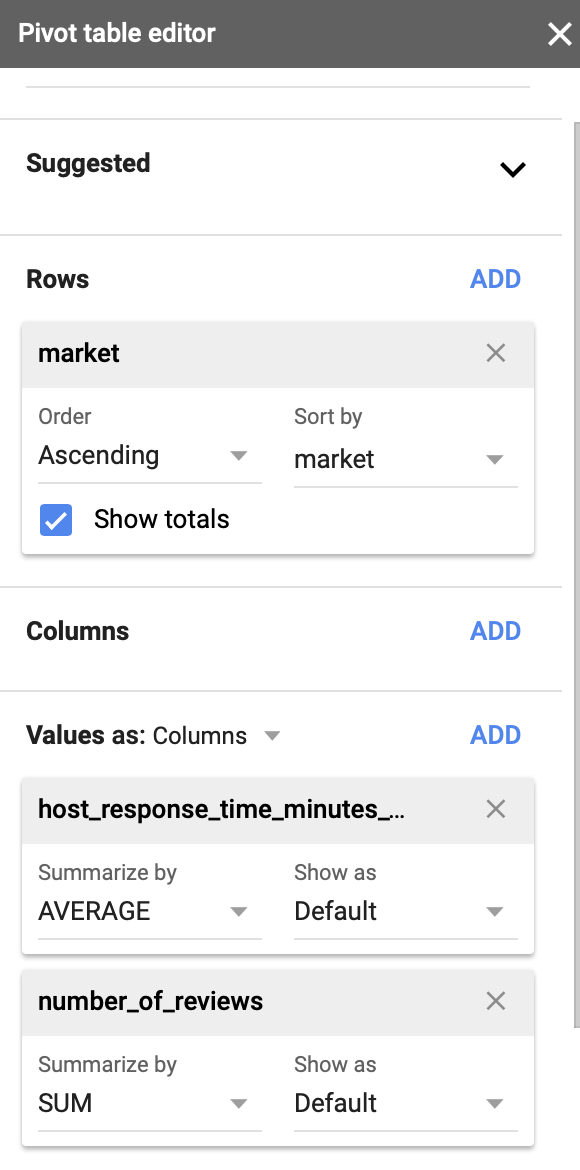
Copy and paste this information in the respective rows in Q1 Market Makeup by Property Type. The percent.total.per.city should auto populate based on your inputs in count.



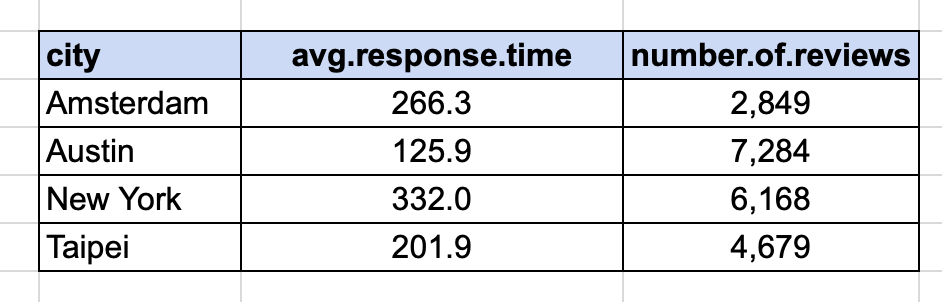
**Correlation Questions**

1. Avg response time and correlation with number of reviews?
2. Based on your analysis, is there an outlier in the dataset?
3. Build out a data visualization to present your findings and share any insights from the data you might have.

Start with the Pivot Table Editor and keep the market variable in the Rows Shelf. In the Values Shelf, add both the host\_response\_time\_minutes\_to\_response / number\_of\_reviews variables, and summarize by Average and SUM, respectively.

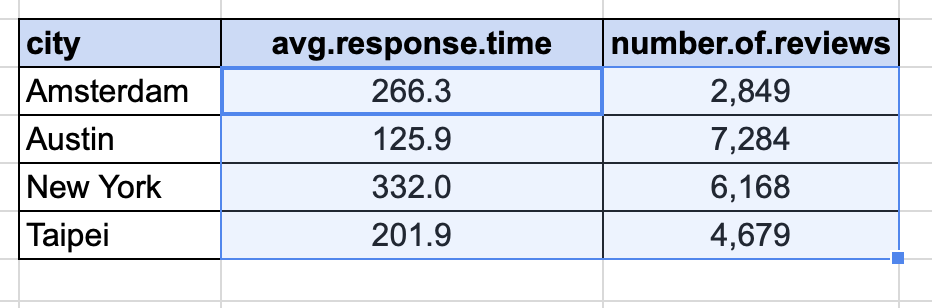


Paste the information you find into the Q2 tab and begin to build a correlation analysis.

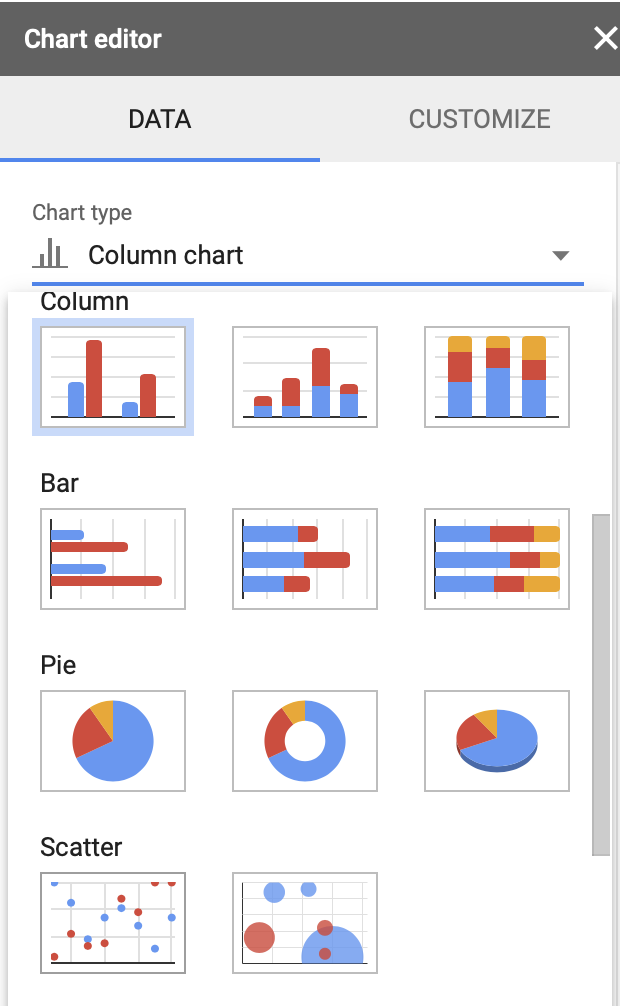


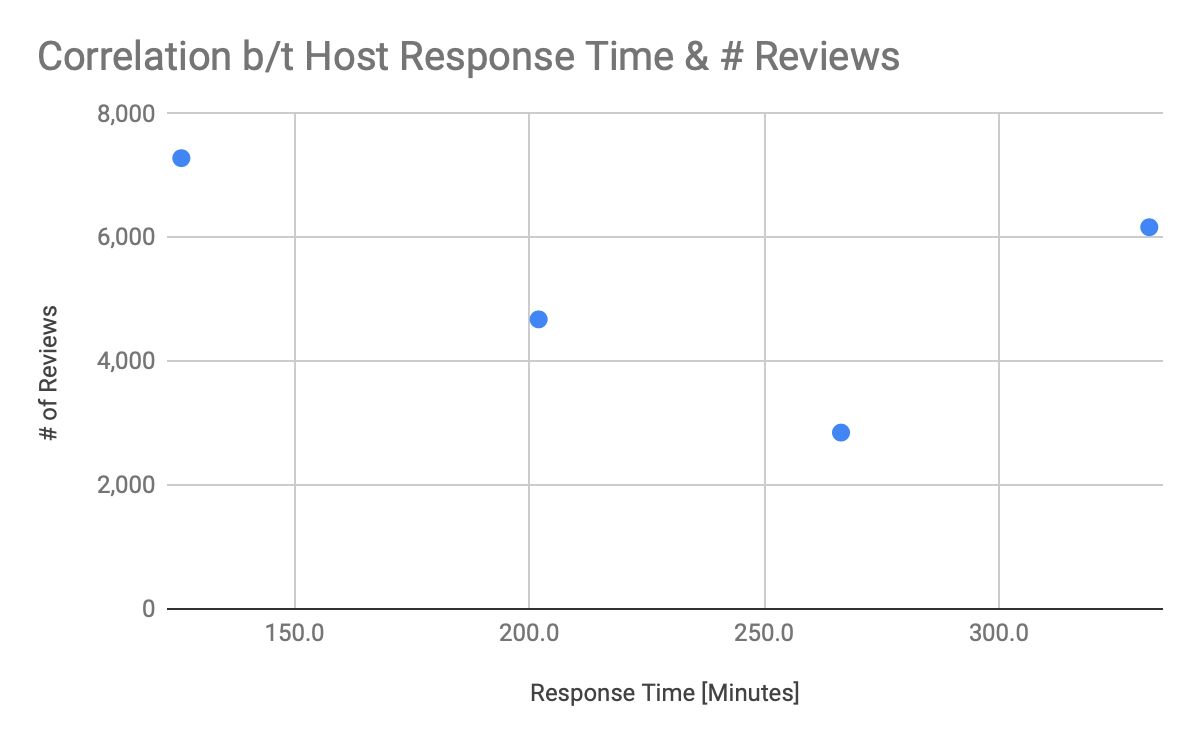
Now, build out a correlation visualization of the data by inserting a chart based on the avg.response.time and number.of.reviews data.

Highlight the data under avg.response.time & number.of.reviews and click Insert -> Chart in the Google Sheets Ribbon.



Click on the Scatter Plot option in the menu of choices for chart types.



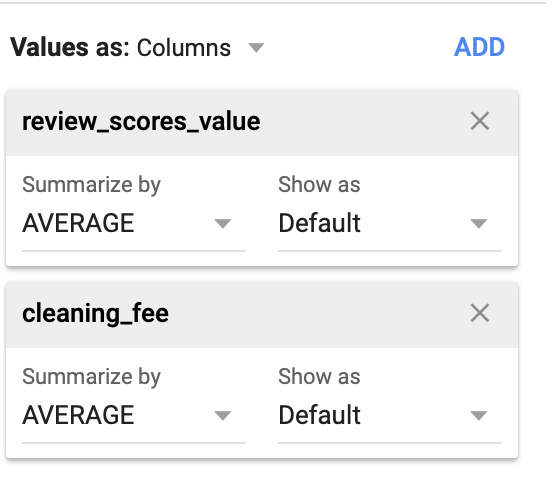


This will populate a correlation scatterplot. As you can see from the data, there is a negative correlation between how long it takes for a host to respond to guests in almost all of the cities that we have analyzed in Airbnb.

**Correlation Questions continued**

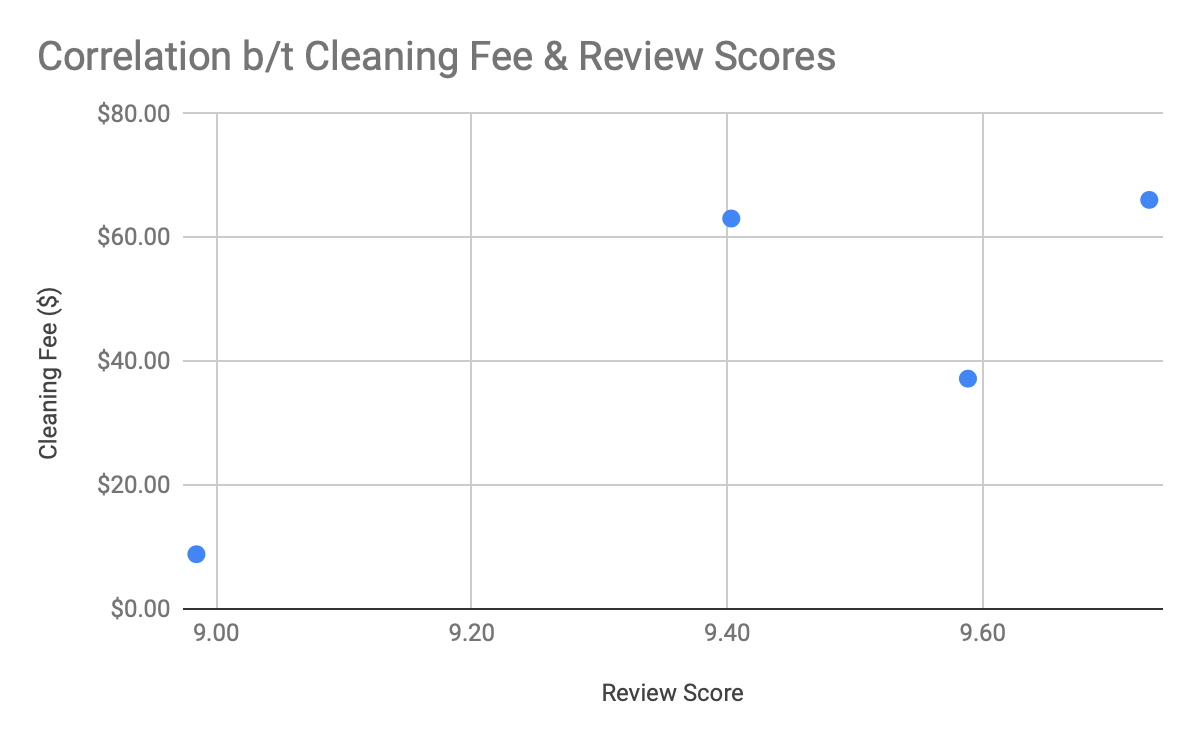
1. Does a higher cleaning fee result in a lower avg rating?
2. Develop a correlation visualization and explain the information that you see.
3. Lend any insights you might have on the matter and develop a visualization to share.

Go back to the Pivot Table Editor and keep the row variable market the same. What we need to do now is add the variables review\_scores\_value & cleaning\_fee to the Values Shelf. Summarize by Average for both of the variables.



As with the previous questions, paste the information from the Pivot Table into tab Q2 and in table titled Correlation b/t Cleaning Fee & Review Score.

Build out a correlation visualization of the data by inserting a chart based on the avg.review.scores.value & avg.cleaning fee. See earlier instructions on how to build out correlation.



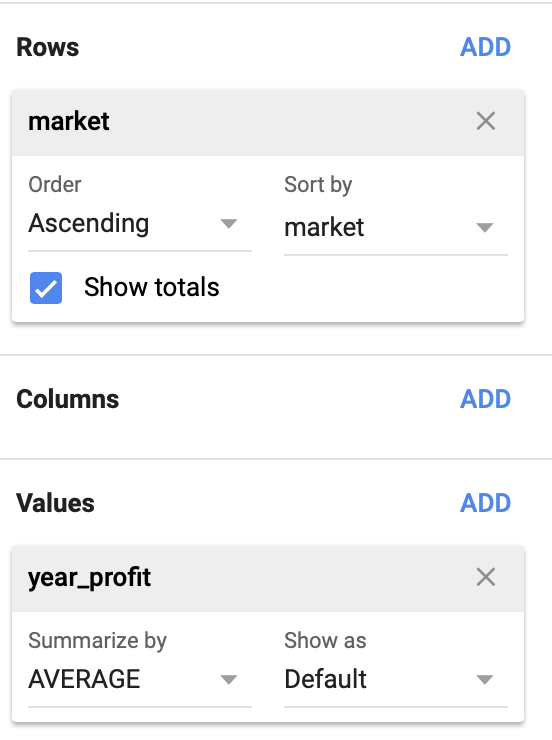
**Modeling Question**

1. Airbnb is concerned with building out a more vibrant host community in its service territory. However, at this point in time, Airbnb does not know what cities are the most profitable for hosts and deserve more of its marketing attention
2. Using the following information/assumptions, identify which cities are the most attractive to new host listings based on yearly profit.
   1. Occupancy Assumptions/ days available per 365 days: Austin 60%;; New York 85%; Amsterdam 65%; Taipei 75% [[occupancy rate analysis](https://www.mashvisor.com/blog/what-airbnb-occupancy-rate-can-you-expect/)]
   2. Assume that revenue is strictly based on nights stayed and price of a room per night.
   3. Hosts pay 3% hosting fee to Airbnb for processing guest payments
   4. Guests also must pay 6% to Airbnb based on the total revenue made per year for guest service fees.
   5. How many listings per city are making more than the average profit per city.
   6. Based off of just this information, what city would you recommend to start with.
   7. Develop a data visualization to present your findings.

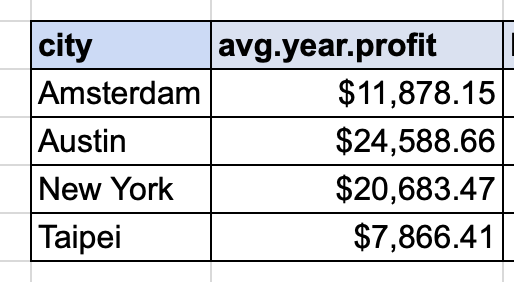
This problem unlike the previous questions has multiple steps to develop before utilizing a Pivot Table. This process involves data munging, or data transformation of the raw\_data set. However, as a business analyst who must joined our team. I will provide you with the equations you must do to answer these questions.

1. Create the year\_revenue variable by multiplying Column W [price\_night] and Column Z
   1. Add a ROUND(Column Z, 0) formula to Column W and then multiply it by Column Z.
2. Create the airbnb\_commission variable by multiplying year\_revenue\_per\_night by 9%, which is the summation of both the hosting fee and guest fees that Airbnb collects found in the questions notes.
3. Then, create a the year\_profit variable by subtracting year\_revenue - airbnb\_commission.

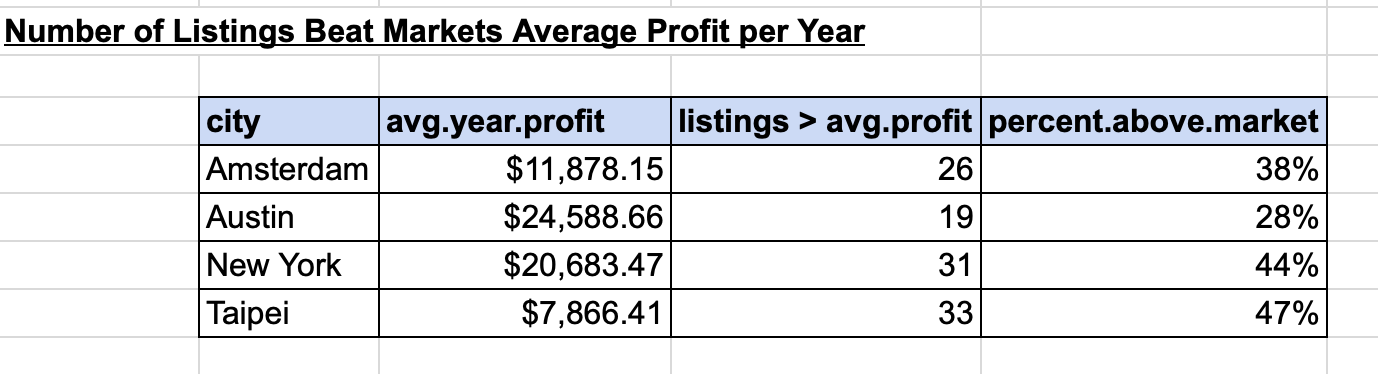
After these 3 variables are created, go back to the Pivot Table Editor and in the Values Shelf, find the average yearly profit for each market.



If done correctly, you will see the following average yearly profits per market.



Paste these numbers into the Q3 table as shown above and the table will auto populate with the insights you need to answer the final question.



Congratulations, I hope that this helped you with your analysis and I look forward to the insights you’ve developed from reading this.