We're going to use a Seattle Airbnb dataset from kaggle, which contains three csv-format datasets: Listings, Reviews, and Calendar. Alex has already combined all three datasets into one: [combined Airbnb data in 2016](). Brief info about each table:

- **Listings**: includes all the possible details about each listing that you can think of, such as geographic location, name, property type & rooms, fees, availability, and reviews.

- **Reviews**: a list of comprehensive reviews of some listings, including its id, review date, reviewer, and comments

- **Calendar:** a full record of each listing's availabilities in terms of its listing id, date, and price

After importing the dataset into Tableau Public, let's prepare the final dataset by joining the Listings and Calendar table together:

- NOTE HERE: we ended up joining 2 tables instead of all three because of the size limitation, let's follow that as well

- First, drag the Listings table to the panel and click the down-arrow button, choose Open

- drag the Calendar table to the panel as well, inner join them by **Listing ID** (this step is important!)

Inner join the Listings and Calendar table by Listings.Id = Calendar.Listing Id

Here's our final table, let's go to worksheet 1 (click bottom left tab named "Sheet 1" and start building our first visualization!

final dataset :)

F or the first visualization we're going to create a bar graph to show the price by each zipcode. This can be helpful because we want to know which geographic locations' prices are more expensive than others, and think about how they can bring more profits to Airbnb hosts.

Let's first drag variable "Zipcode" to the "Columns" bar and exclude the Null value. Next drag the measure value "Price" to the "Rows" and select "Average" as the calculated formula.
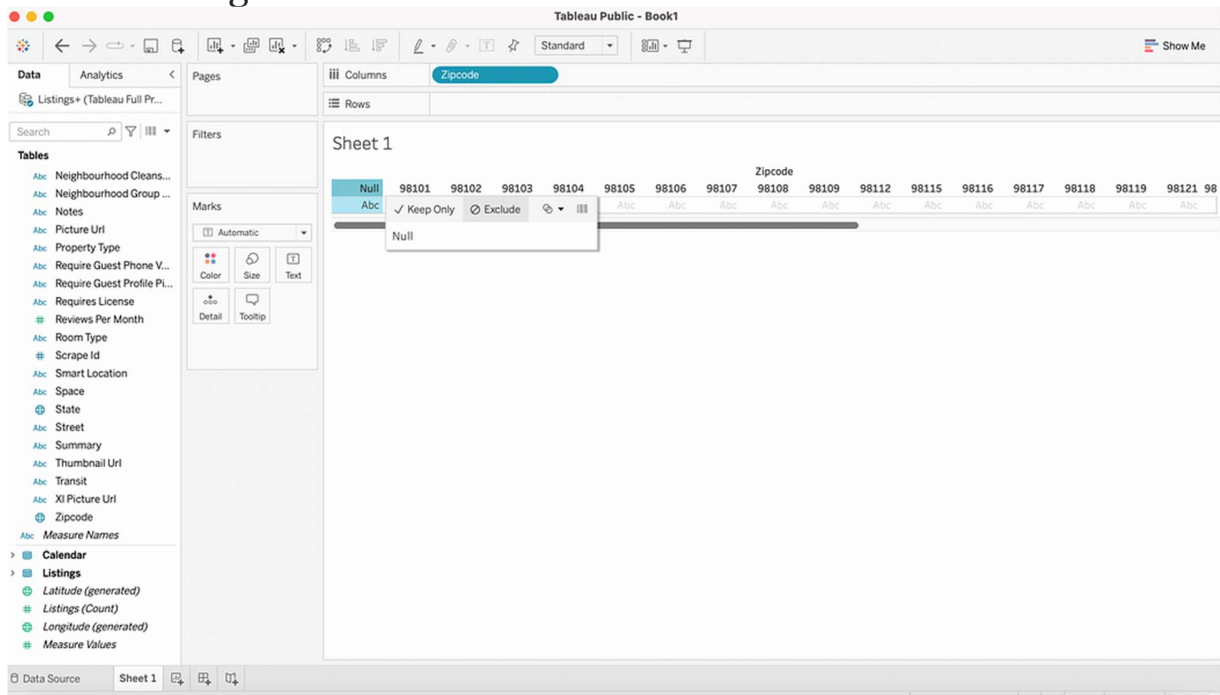
exclude the Null value
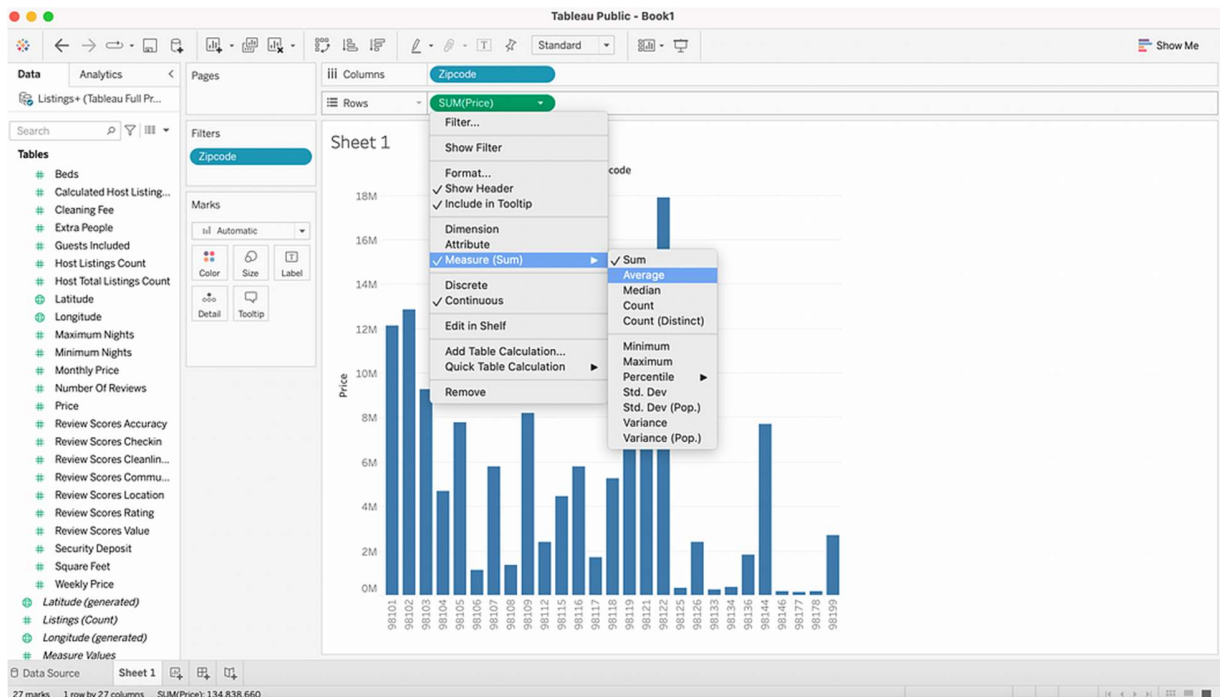
calculated by Average instead of SUM

Now let's sort the results in descending order and give this graph some color — based on zipcode. We need to drag the variable "Zipcode" to the "Marks" panel, and change it from "Detail" (the triangle dots) to show in "Color" (rectangle colorful dots). Remember to change the graph title too, let's give it "Price By Zipcode". And woohoo — here's our first visualization, doesn't it already look so much better?!
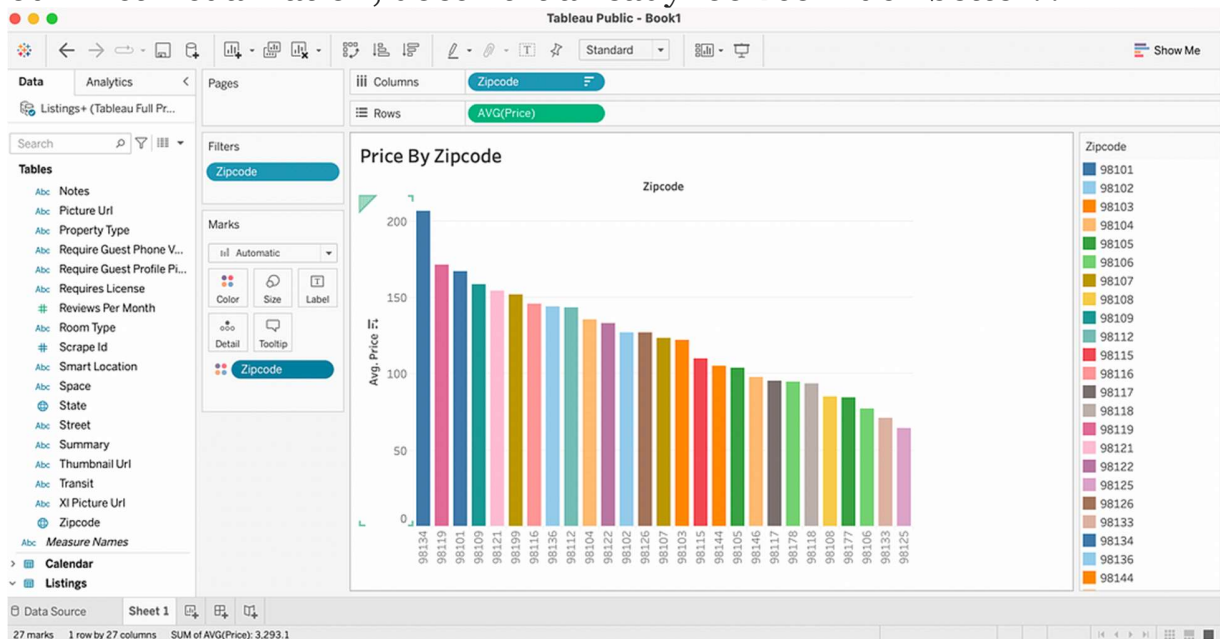
first visualization :)

For the second visualization, we will be creating a map to illustrate price by zipcode as well. Let's go to the second sheet, drag dimension "Zipcode" to the column bar, and select the map in "Show Me" list.
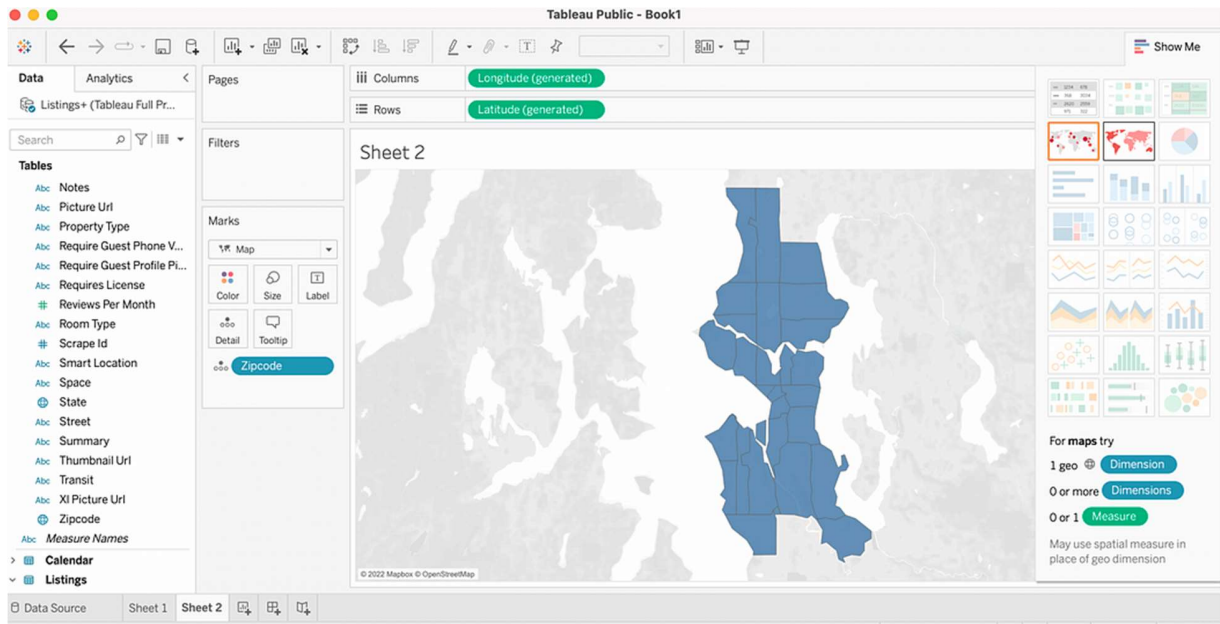
We now have a map of all the Zipcodes of Airbnbs! Let's use color and labels to differentiate the locations. To achieve this, we will drag the dimension "Zipcode" to the Marks panel, change the display from "detail" to "color", and change the other one from "detail" to "label".
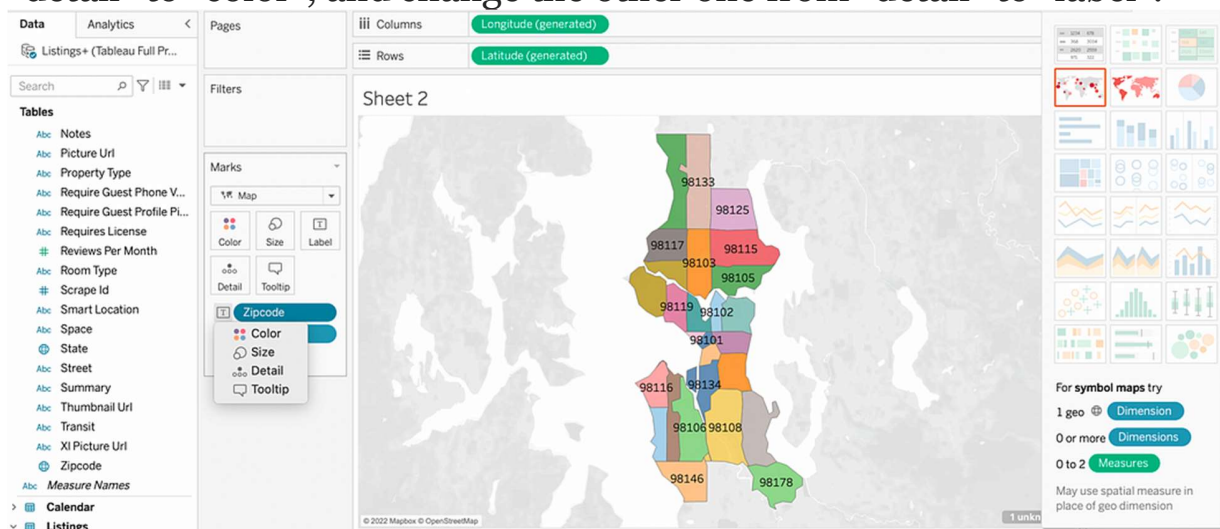
In order to know the average airbnb price by location, we can display the price as label as well. Drag the measure "Price" to Marks panel,

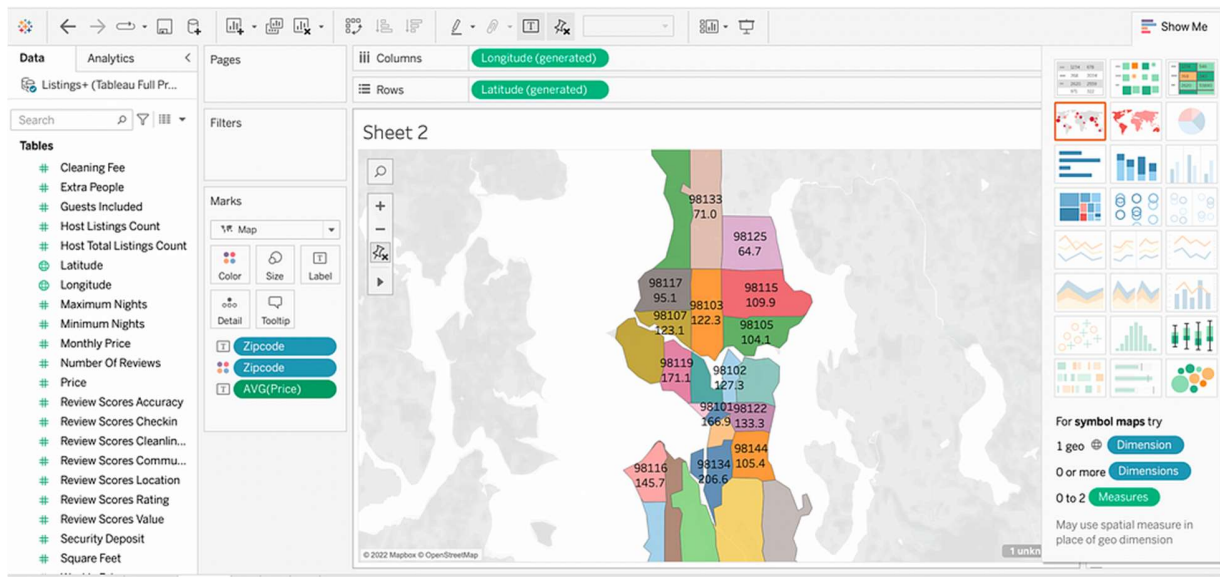change the calculated formula to "AVG", and then display it as "label". And here's our second visualization, which correlates with the first bar graph, but just in a more visual way. Now let's go to the third visualization!


Besides getting familiar with the airbnb price by different locations, we might also be interested in when does the airbnb price hike up a lot, or when do people spend the most in airbnb. We can use the "Calendar" table to help us understand this and build a time series graph. First, let's drag the dimension "Date" to the columns bar, and change the display from "Year" to "Week".
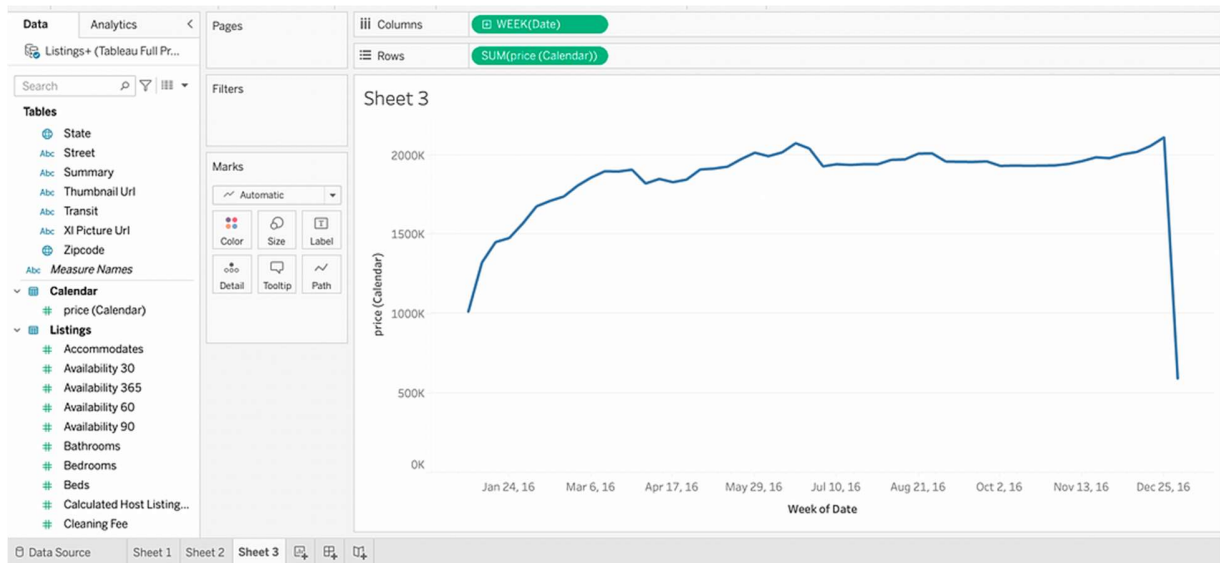
We can see there's a huge drop at the end of December, 2016 because of the lack of data in 2017. Let's filter 2017 out by adding a filter as well - click the little "down" arrow from "WEEK(Date)" and select filter, and drag the date back to 12/31/2016.
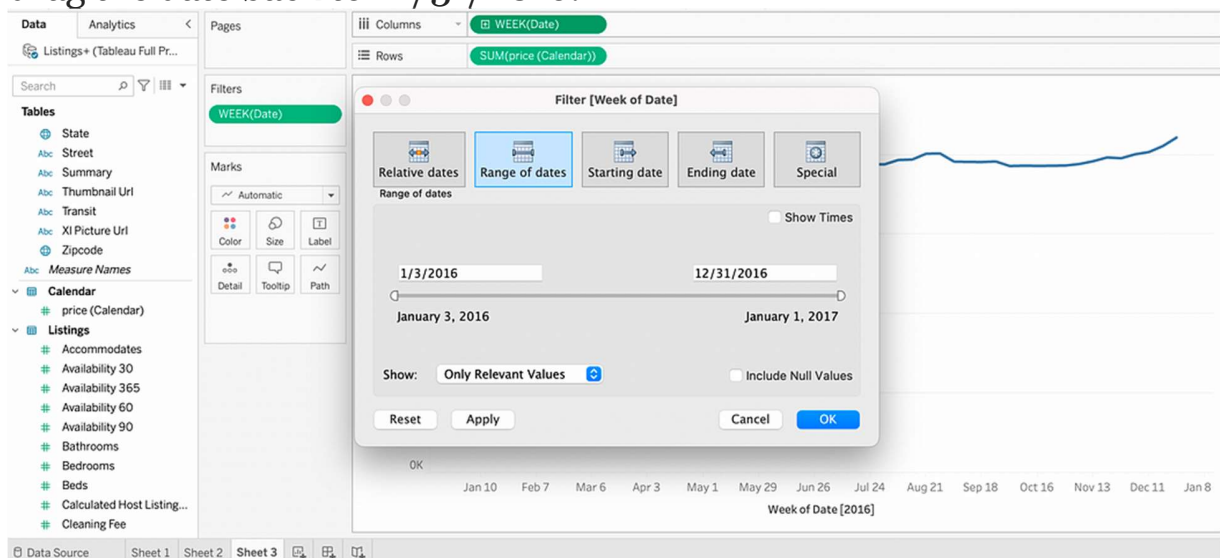
To show the minimum & maximum revenue by each week in 2016, we can drag the measure "Price" to the Marks panel, change display to

"label", click the "Label" box and select "Min/Max" from the "Marks to Label" section.
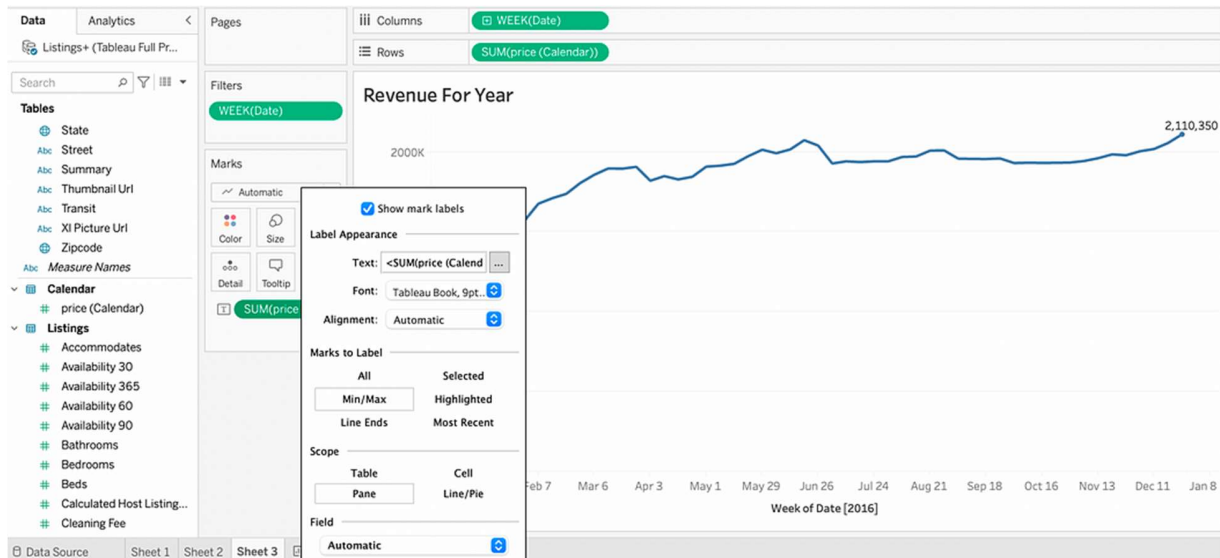
And here's our final line chart! It's kinda obvious that by the end of 2016, airbnb revenue reaches its peak because of the holiday season. Ok we have 2 more visualizations to go, let's move on!
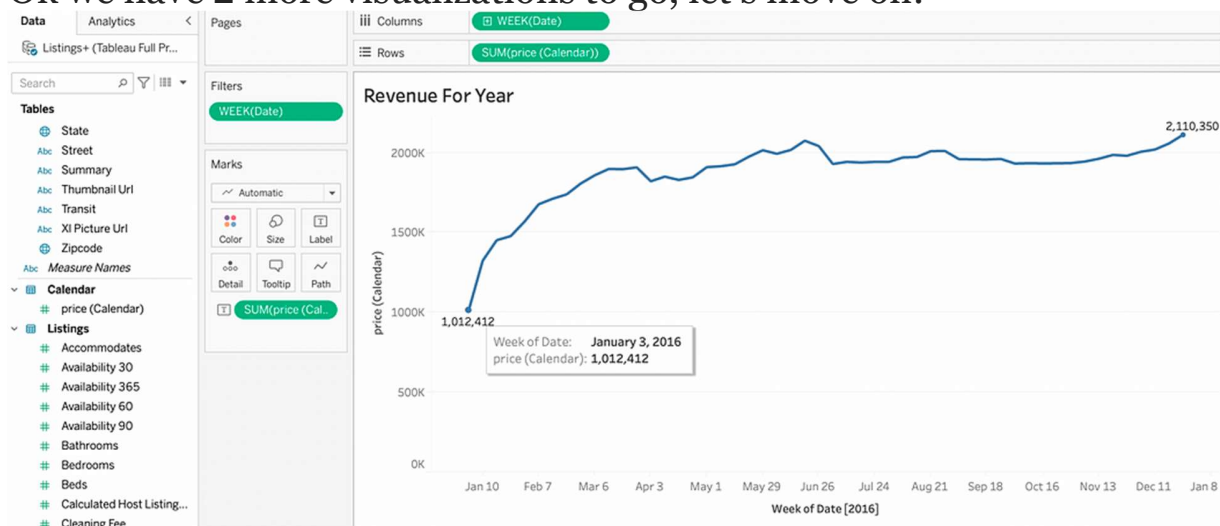

So far we have looked at the airbnb price by different locations & its revenue over the year, now let's check out the price difference by the #

of rooms in different airbnbs. To do this, let's create a graph with the number of bedrooms as the column, and average of price as row. Since "bedroom" is a measured value, we need to first convert it to a "Dimension", and then drag the new variable to the Columns bar. Now drag the "Price" value to the Rows bar, and change the formula from "SUM" to "AVG".
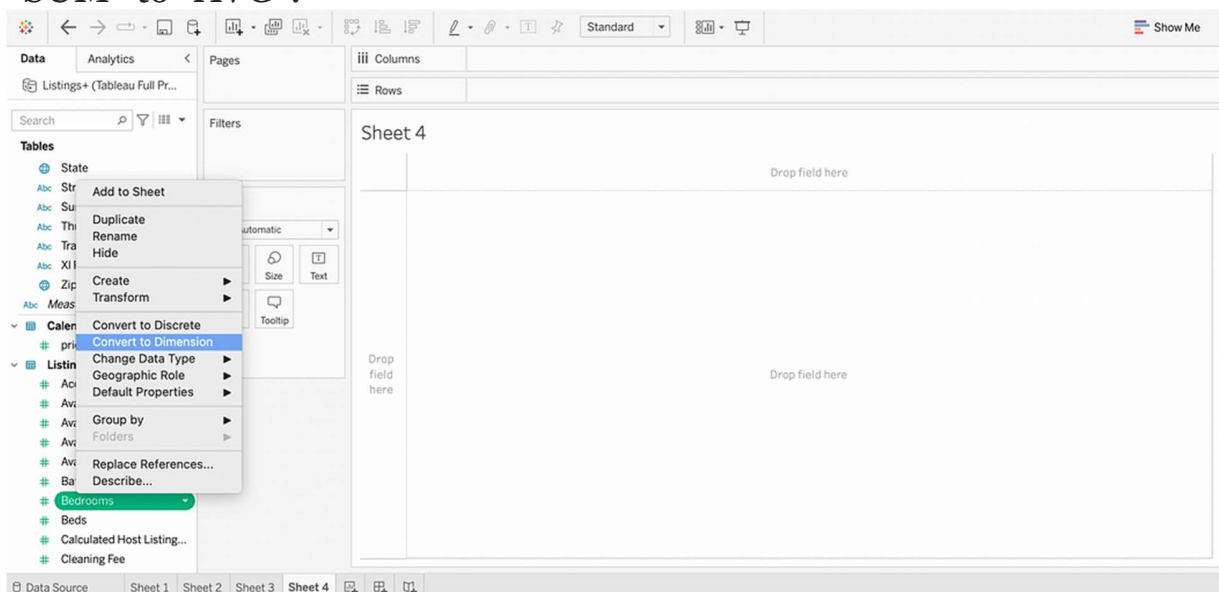

We also added a label to illustrate the price for airbnbs with different numbers of rooms. We can clearly see that the more number of rooms, the higher the airbnb price goes.
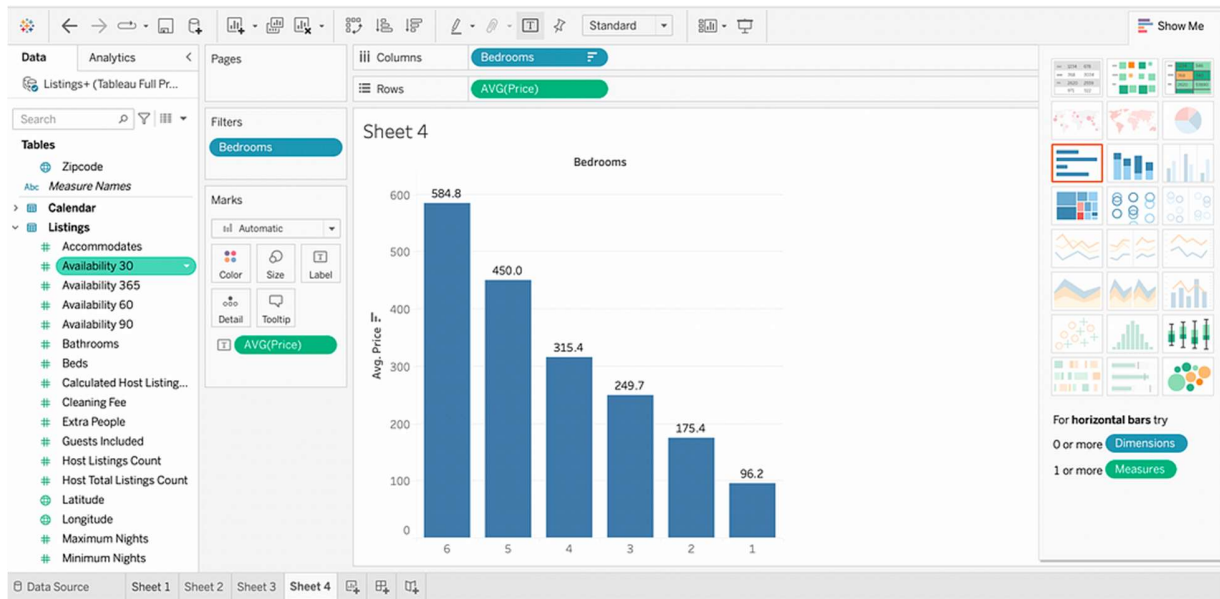
Inthe last visualization, let's just create a simple table to count the number of listings for each type of airbnb. Let's first drag the dimension "Bedrooms" to the Rows bar, and put the "Id" variable to the Marks panel. Change the display from "detail" to "Text", and switch the formula to "Count (Distinct)" since we want to know how many listings are there for each type of airbnb.

And here's our final table — discount count of listings by different # of bedrooms. As the number of bedrooms goes up, there are fewer listings that's available, and there're so many options if we're just looking for one bedrooms!

It's time to put everything together and create our dashboard! This is a super easy step, just drag each sheet to the dashboard panel and order it in whatever format you like :) You can save and share this dashboard to Tableau Public, here's my link: [Airbnb Pricing](#)
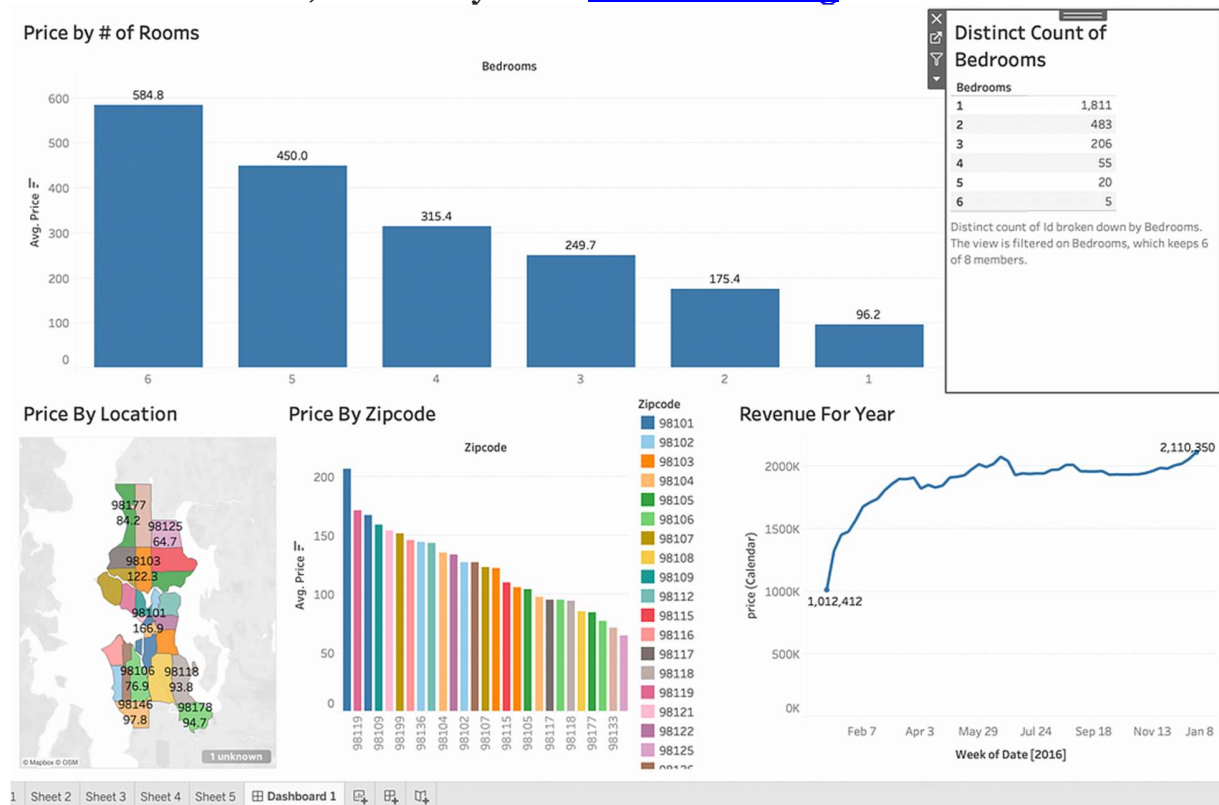
Thanks so much for reading, hope you have a nice day!