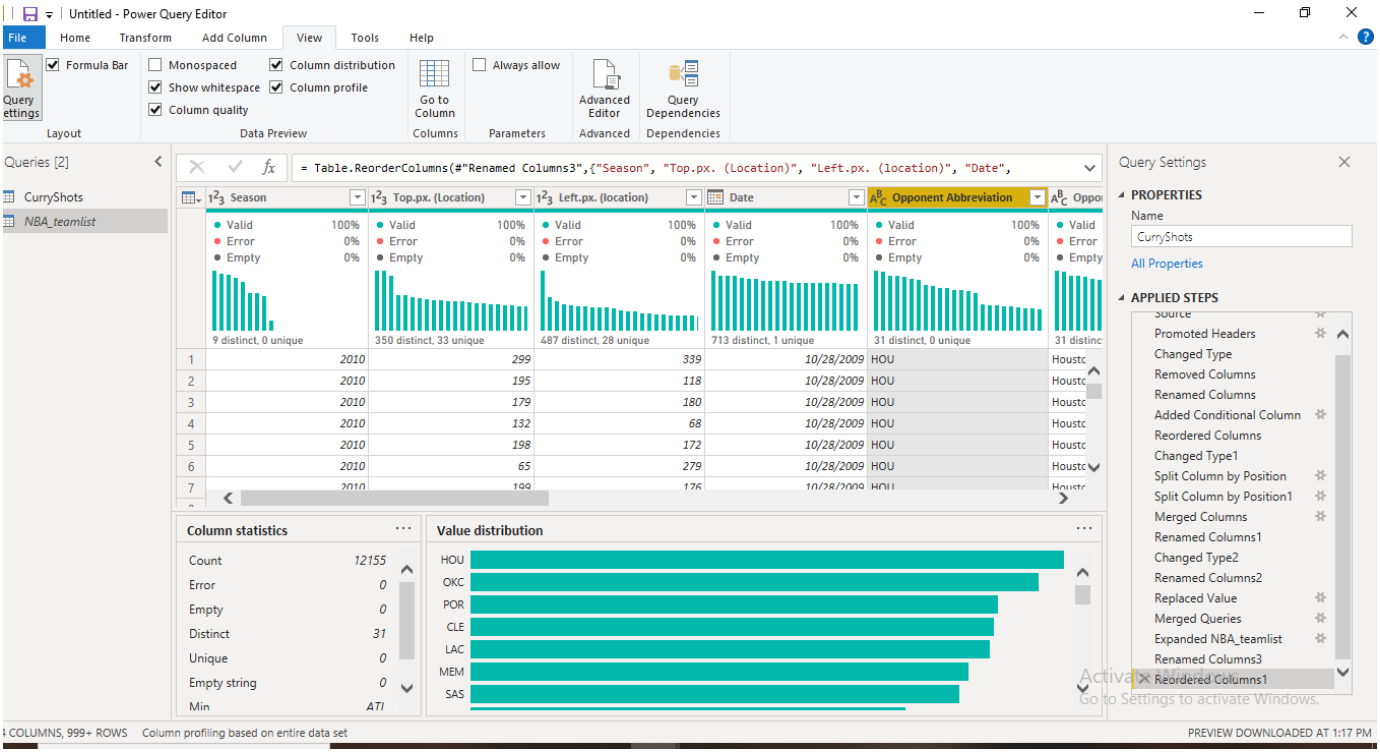
WEEK 2

Merge Queries & Column Profiling

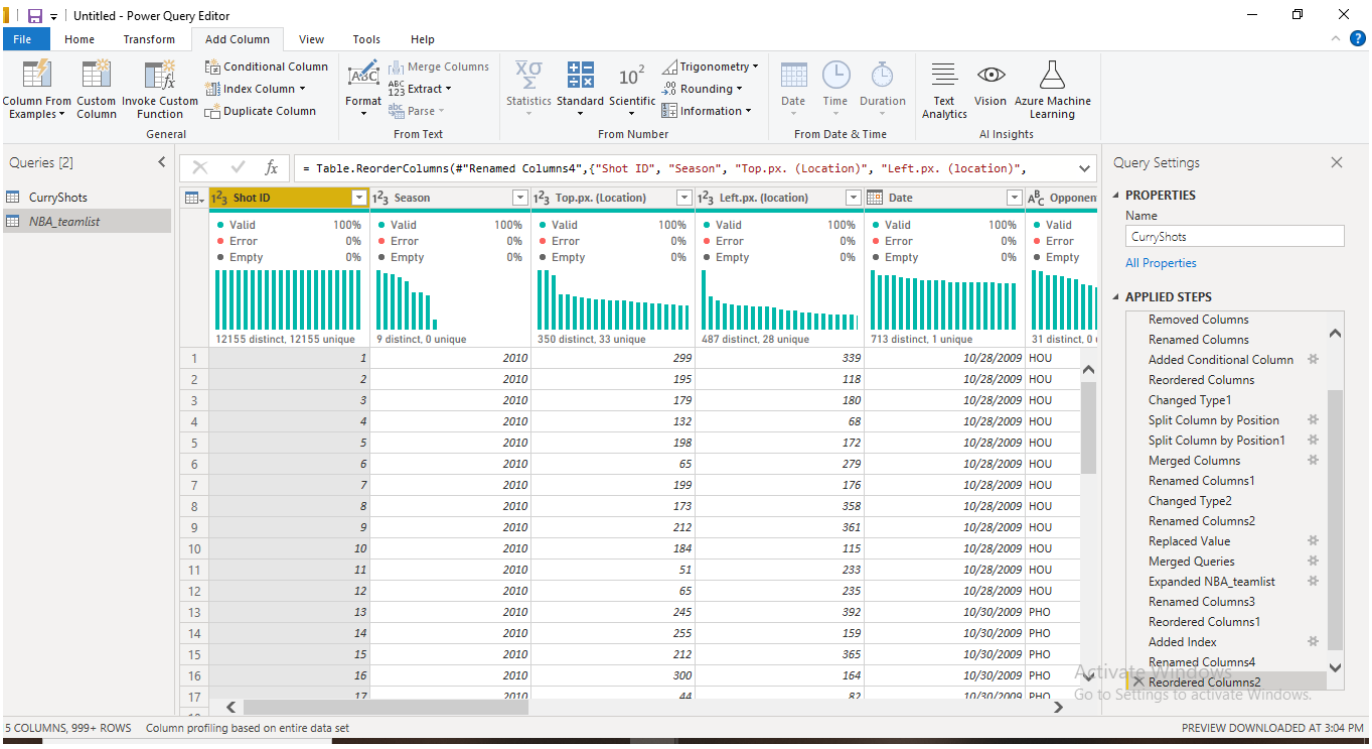
We can make a lot of transformations in our data while in Power Query to make it clean before loading.

Here, we explored a feature in Power Query that allows us to combine data from two sources into one to create a joined table. Unfortunately, when merging, Power Query could not match a few contents, leading to a messy dataset. We tracked down those empty rows via column profiling. Doing so, we found out that our dataset was mistakenly coded. That is why Power Query was not able to match correctly. Therefore, we replaced the mistaken values with the right ones. On replacing, Power Query found the correct matches throughout the dataset.

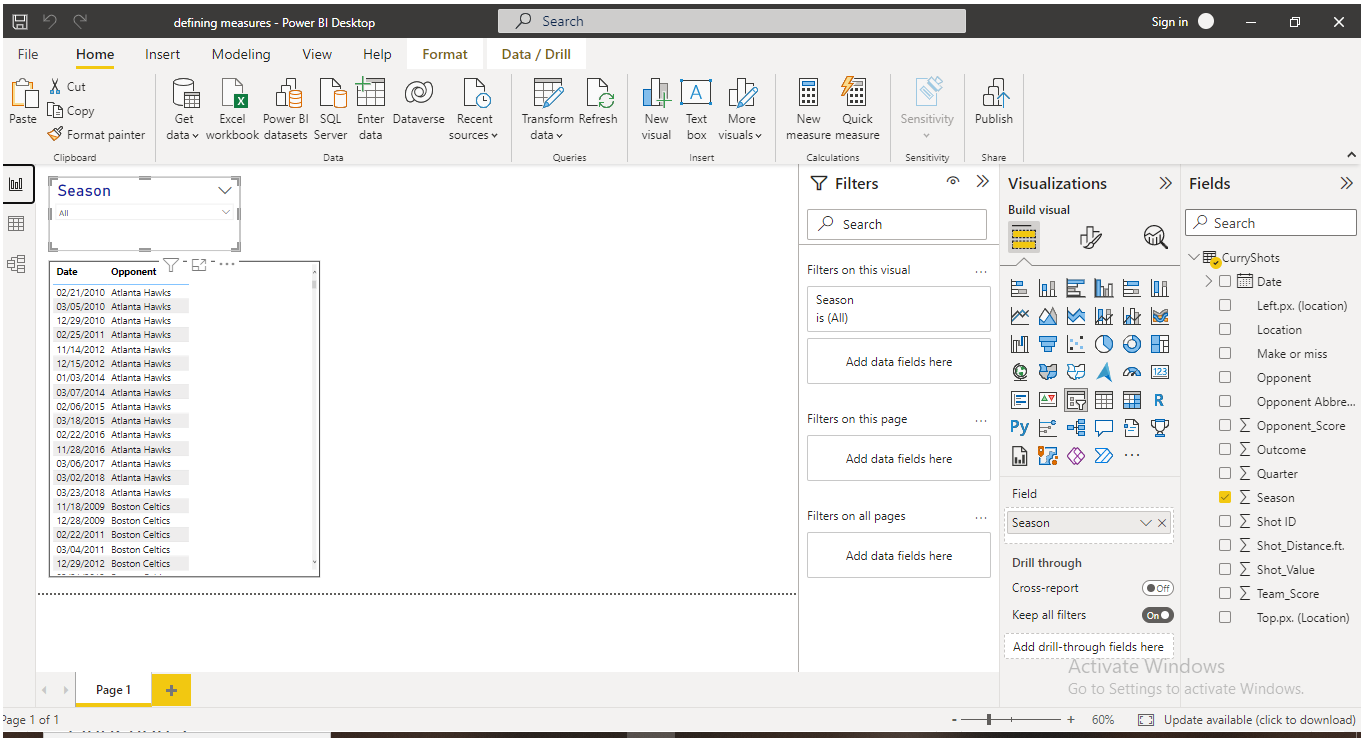
We also turned off the enable load option for our second data source to prevent it from being loaded in our Power BI model and to keep the Power BI desktop streamlined.



One last thing I did before loading was to add an index column named Shot ID.



Next, since there are around 12k shots that Curry took, we will add a slicer in our report and a season column. I also created a table column and added the date and opponent columns. I also changed the formatting of the date from date hierarchy to simply date. It saves space in the report.



Now, I created a new explicit measure in our model that we will reuse further: Total Field Goals (FG) Attempted.

```
Total FG Attemped = DISTINCTCOUNT(CurryShots[Shot ID]
```

Our next measure is Total FG Made which returns the values of all the shots Curry made, whose outcome is equal to 1. For this, we inserted CALCULATE function. Our formula was as below:

```
Total FG Made = CALCULATE( [Total FG Attemped], CurryShots[Outcome] = 1 )
```

To calculate Total Field Goal (FG) Percentage, we used the following formula:

```
Total FG Percentage = DIVIDE ([Total FG Made], [Total FG Attemped], 0)
```

To calculate 3points Field Goals Attempted, I created another measure:

```
3pt FG Attemped = CALCULATE ( [Total FG Attemped], CurryShots[Shot_Value] = 3)
```

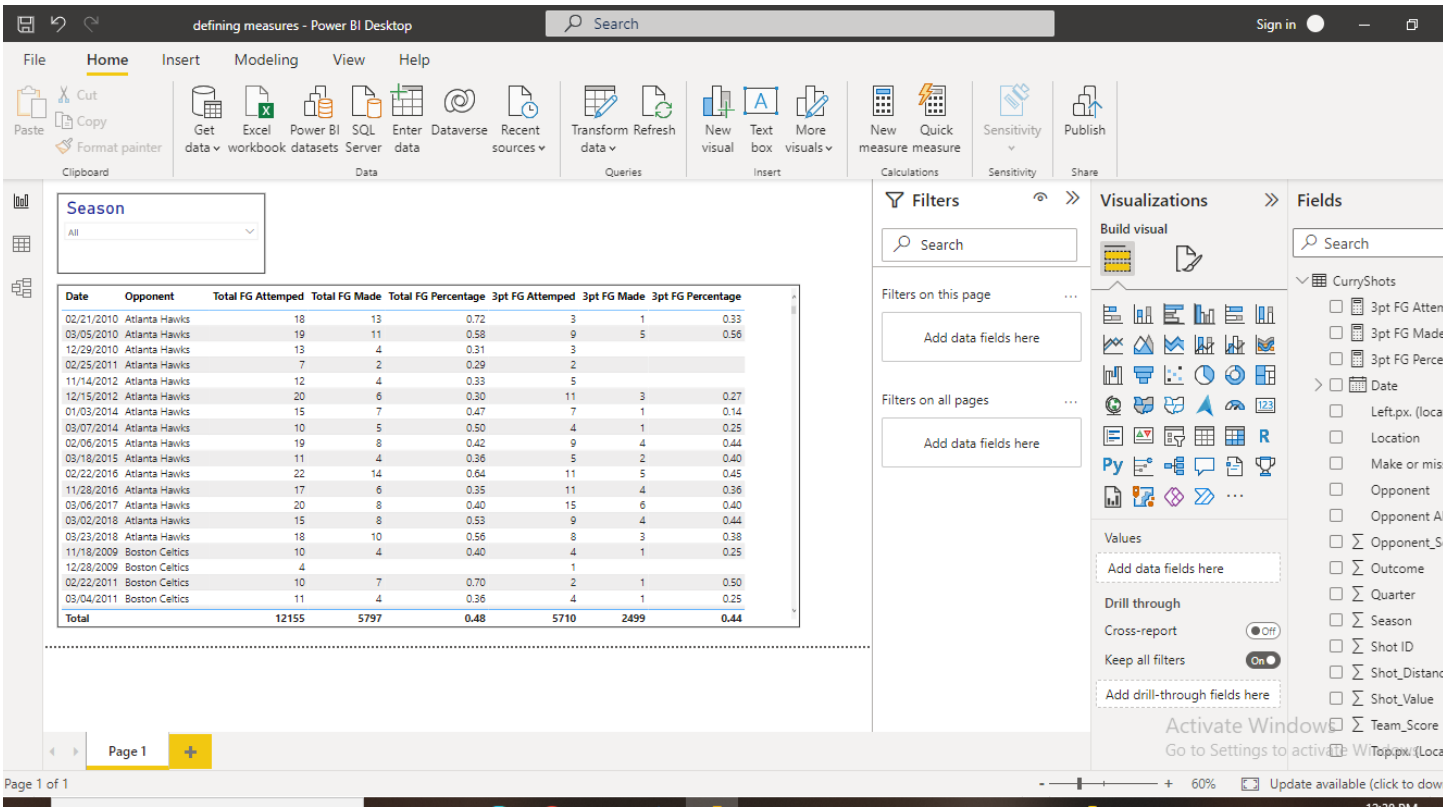
For 3 points Field Goals Made, I created the new measure with the formula:

```
3pt FG Made = CALCULATE([Total FG Made], CurryShots[Shot_Value] = 3)
```

Finally, to calculate the 3pt FG Percentage, we created our sixth explicit measure using the formula:

```
3pt FG Percentage = DIVIDE ([3pt FG Made],[3pt FG Attemped], 0)
```

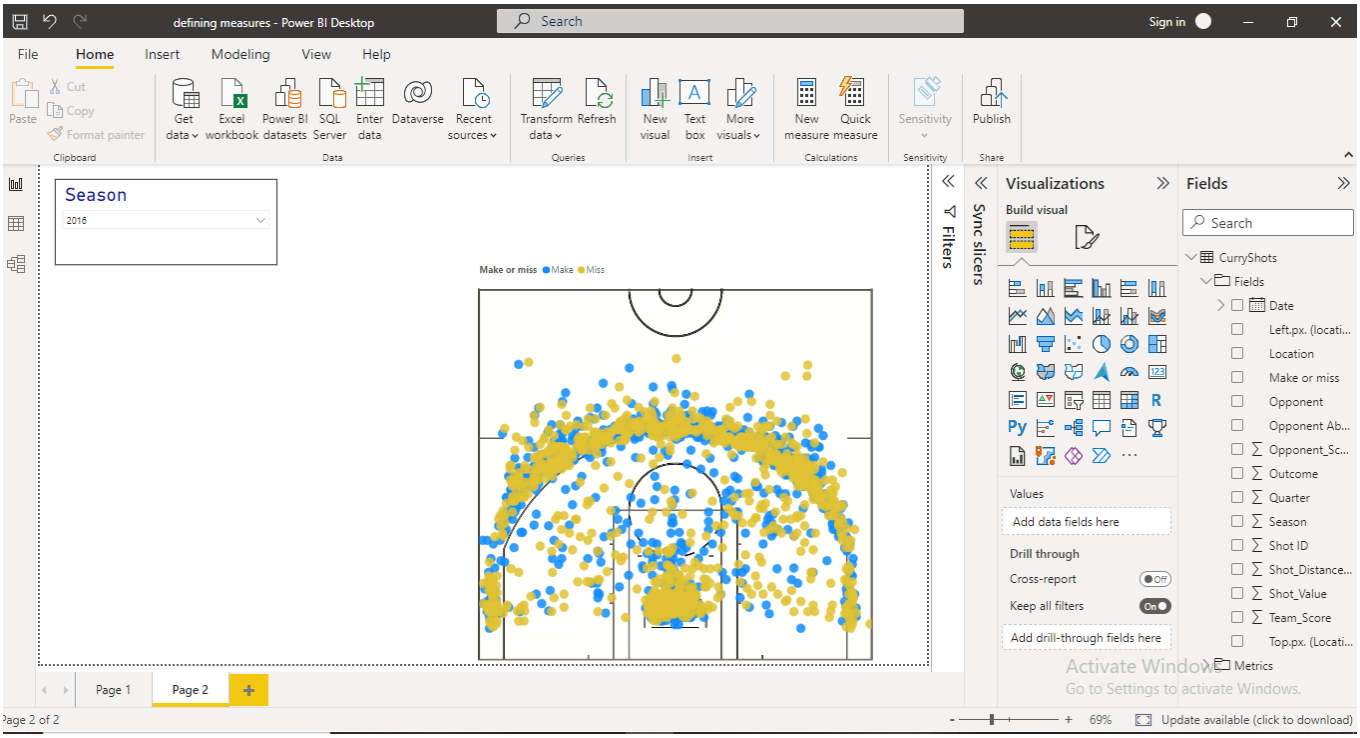
Next, we will go back to our report page, and in the table visual we created earlier, we will add all these six explicit measures.



Date	Opponent	Total FG Attemped	Total FG Made	Total FG Percentage	3pt FG Attemped	3pt FG Made	3pt FG Percentage
02/21/2010	Atlanta Hawks	18	13	0.72	3	1	0.33
03/05/2010	Atlanta Hawks	19	11	0.58	9	5	0.56
12/29/2010	Atlanta Hawks	13	4	0.31	3		
02/25/2011	Atlanta Hawks	7	2	0.29	2		
11/14/2012	Atlanta Hawks	12	4	0.33	5		
12/15/2012	Atlanta Hawks	20	6	0.30	11	3	0.27
01/03/2014	Atlanta Hawks	15	7	0.47	7	1	0.14
03/07/2014	Atlanta Hawks	10	5	0.50	4	1	0.25
02/06/2015	Atlanta Hawks	19	8	0.42	9	4	0.44
03/18/2015	Atlanta Hawks	11	4	0.36	5	2	0.40
02/22/2016	Atlanta Hawks	22	14	0.64	11	5	0.45
11/28/2016	Atlanta Hawks	17	6	0.35	11	4	0.36
03/06/2017	Atlanta Hawks	20	8	0.40	15	6	0.40
03/02/2018	Atlanta Hawks	15	8	0.53	9	4	0.44
03/23/2018	Atlanta Hawks	18	10	0.56	8	3	0.38
11/18/2009	Boston Celtics	10	4	0.40	4	1	0.25
12/28/2009	Boston Celtics	4			1		
02/22/2011	Boston Celtics	10	7	0.70	2	1	0.50
03/04/2011	Boston Celtics	11	4	0.36	4	1	0.25
Total		**12155**	**5797**	**0.48**	**5710**	**2499**	**0.44**

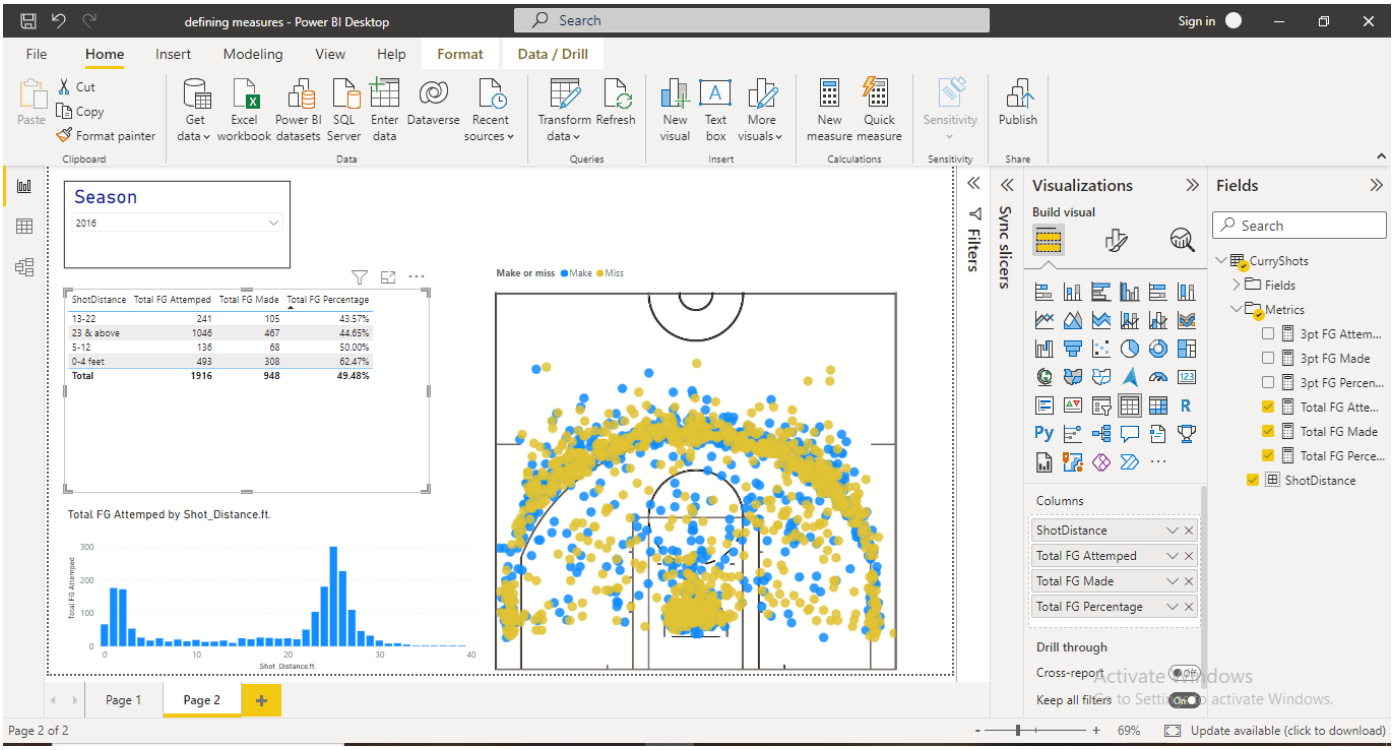
Sync Slicers & Scatter Charts

Here, we first synced slicers. Next, we plotted our shot data (make or miss) on top of the image of a basketball court image using a scatter chart.



Grouping, Card Visuals, and Conditional Formatting

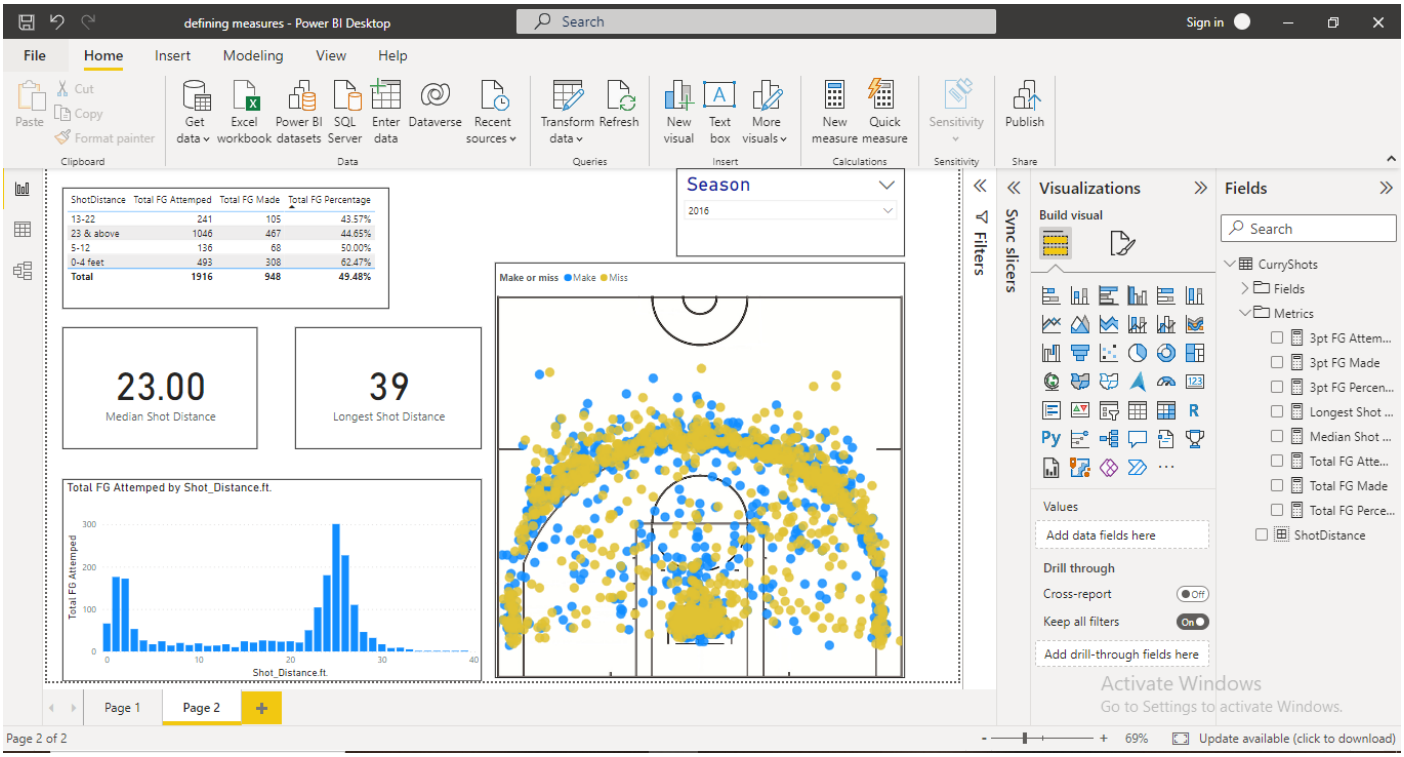
Here, we grouped the field: shot_distance and created a table demonstrating these groups and the total FG attempted, made, and percentage.



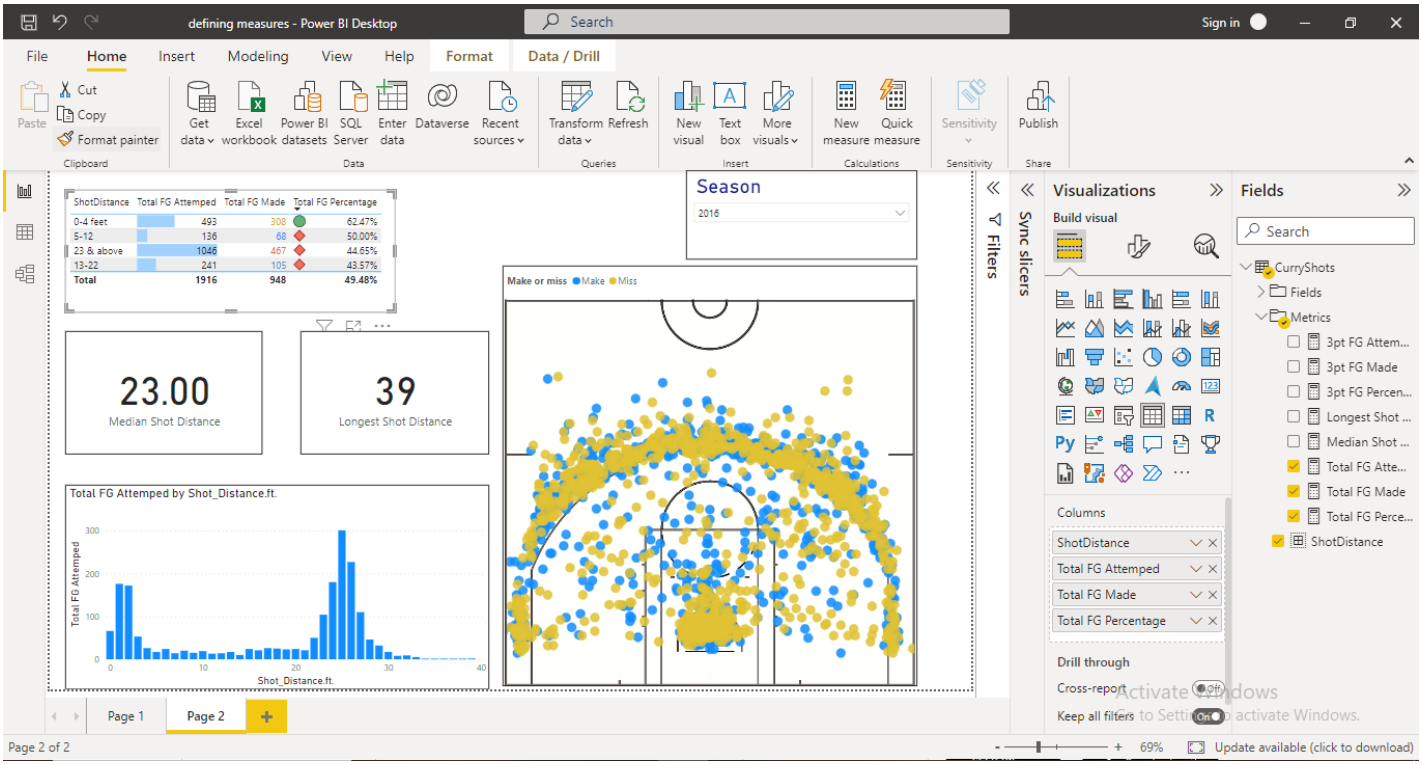
In this case, we wanted to determine Curry's longest and median shot distance, so we created cards. For this, I created two explicit measures in our data model.

For median: `Median Shot Distance = MEDIAN(CurryShots[Shot_Distance.ft.])`

For longest Shot Distance = `Longest Shot Distance = MAX(CurryShots[Shot_Distance.ft.])`



Finally, we did some conditional formatting on the table we created earlier.


defining measures - Power BI Desktop
Search
Sign in
File
Home
Insert
Modeling
View
Help
Format
Data / Drill
Cut
Copy
Format painter
Paste
Get data
Excel workbook
Power BI datasets
SQL Server
Enter data
Dataverse
Recent sources
Transform data
Refresh
New visual
Text box
More visuals
New measure
Quick measure
Sensitivity
Publish
Clipboard
Data
Queries
Insert
Calculations
Sensitivity
Share
ShotDistance
Total FG Attemped
Total FG Made
Total FG Percentage
0-4 feet 493 308 62.47%
5-12 136 68 50.00%
23 & above 1046 467 44.65%
13-22 241 105 43.57%
Total 1916 948 49.48%
Season
2016
23.00
Median Shot Distance
39
Longest Shot Distance
Make or miss
Make
Miss
Total FG Attemped by Shot_Distance.ft.
Total FG Attemped
Shot_Distance.ft.
Sync slicers
Filters
Visualizations
Build visual
Columns
ShotDistance
Total FG Attemped
Total FG Made
Total FG Percentage
Drill through
Cross-report
Keep all filters
Fields
Search
CurryShots
Fields
Metrics
3pt FG Attem...
3pt FG Made
3pt FG Percen...
Longest Shot ...
Median Shot ...
Total FG Atte...
Total FG Made
Total FG Perce...
ShotDistance
Page 1
Page 2
Page 2 of 2
69%
Update available (click to download)