MIS0855: Data Science
In-Class Exercise for Mar 30 - Apr 1 – Working with "Pivot Tables" in Tableau

Objective: Work with dimensional data to navigate a data set

Learning Outcomes:

- Summarize a table of data organized along dimensions
- Create hierarchies to enable drill-up/drill-down capability
- Select the correct dimensions and measures to answer a question

This exercise takes you through the steps to create "pivot-style" tables in Tableau. I call them "pivot-style" tables because Tableau doesn't use that terminology. However, the functionality is very similar to the Pivot Table feature you see in Microsoft Excel. Therefore, most things you will do in this exercise you can also do in those programs. However, in Tabelau, it's a little easier to create the tables and a lot easier to turn the tables into charts!

You'll be working with data from a site called Spreadsheet Sports (http://www.spreadsheet-sports.com). This data set contains player statistics from the 2013-2014 NCAA Basketball season. It includes player positions, their division and scoring data. You'll find a "Data Dictionary" tab that explains each field.

Part 1: Download the spreadsheet and open it in Tableau

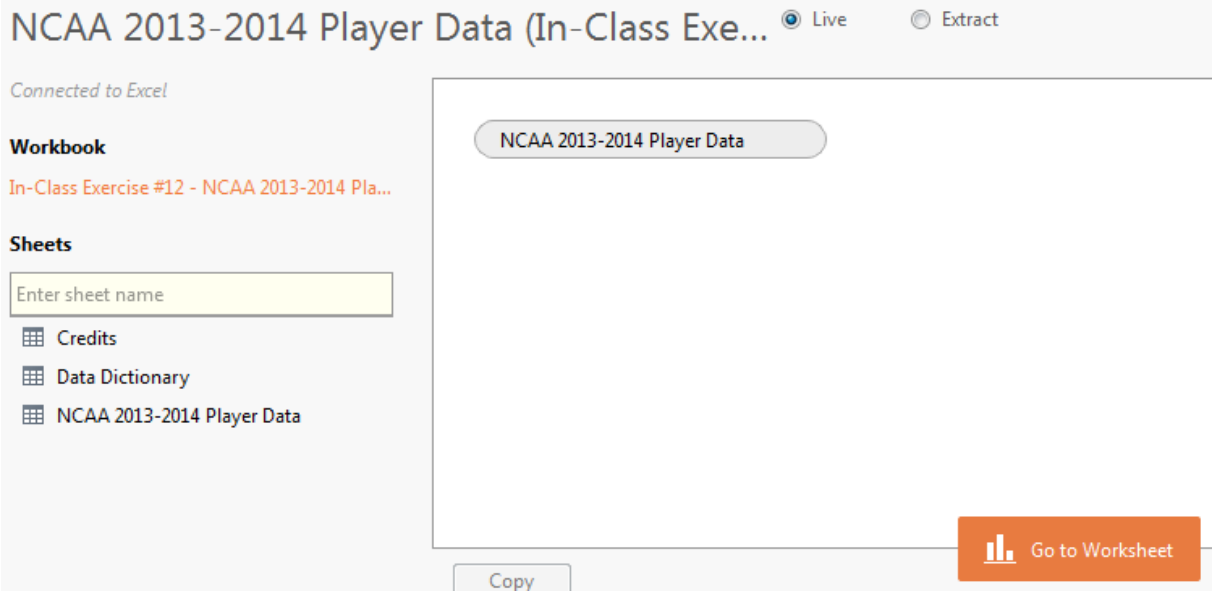
1) Download "NCAA 2013-2014 Player Stats.xlsx." Save the file to your computer.

2) Open the file in Excel and take a quick look through it. Spend a couple minutes in looking at the Data Dictionary.

3) Start Tableau, click "Connect to Data" and then "Microsoft Excel."

4) Open the player statistics file.

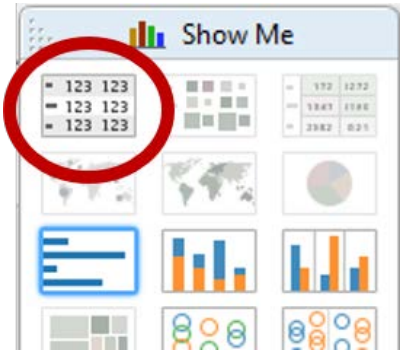
5) Drag the “NCAA 2013-2014 Player Data” sheet to the blank workspace area. Wait until the data shows up and click “Go to Worksheet.”



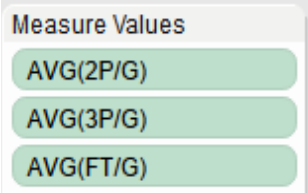
Part 2: Analyze Aggregate Player Statistics by School

1) Drag the “School” dimension to the Rows shelf.

2) Drag 2p/G, 3p/G, and Ft/G to the Columns shelf. The chart may change to a bar chart. If it does, change it back to a Text table by clicking that icon in the “Show Me” area.



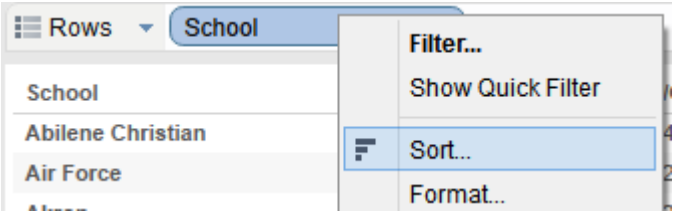
3) Change the measure value calculations from SUM to AVG by right-clicking on each one and selecting Measure/Average. You're now looking at the average player per-game scoring statistics, organized by school.



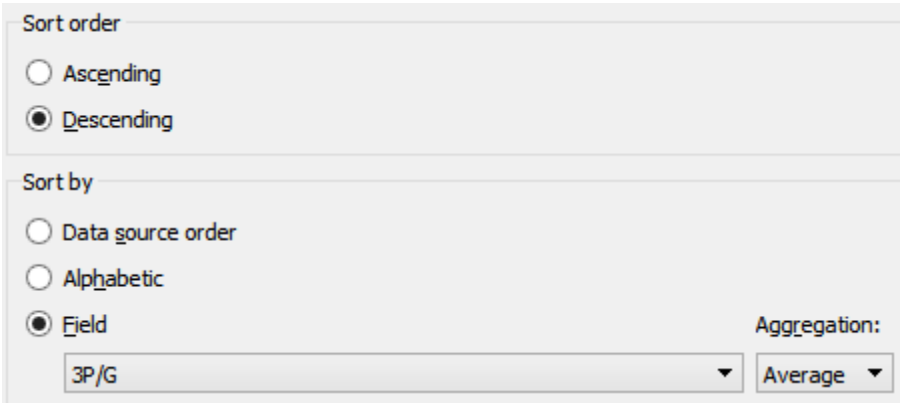
For example, each player at Abeline Christian scores an average of 1.577 2-point shots per game and 0.763 3-point shots per game. This makes sense – 3-point shots are more difficult to make, so you'd expect there to be less of them.

School	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
Abilene Christian	1.577	0.763	1.646
Air Force	1.118	0.560	1.027
Akron	1.577	0.723	1.328
Alabama	2.052	0.640	1.817
Alabama A&M	1.207	0.463	1.148
Alabama State	2.062	0.601	1.612
Alabama-Birmingham	1.964	0.507	1.549
Albany (NY)	1.566	0.477	1.474

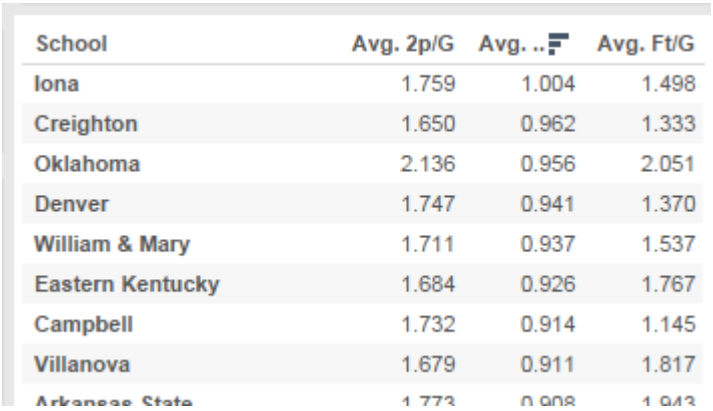
4) We want to know which school has the best 3-point shot scoring average among its players. To do this, we can sort the table. Click on "School" in the Rows shelf and select "Sort…"



5) Choose “Descending” for Sort order, “Field” for Sort by, and the select the “3p/G” field and “Average” for aggregation. Click “OK.”



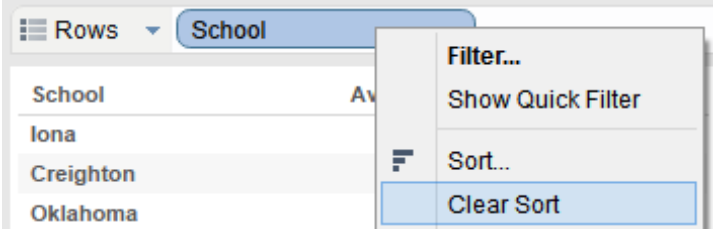
6) You now see an icon indicating that this column is sorted in a descending order and that Iona’s players score, on average, 1.004 3-point shots per game.



School	Avg. 2p/G	Avg. ..	Avg. Ft/G
Iona	1.759	1.004	1.498
Creighton	1.650	0.962	1.333
Oklahoma	2.136	0.956	2.051
Denver	1.747	0.941	1.370
William & Mary	1.711	0.937	1.537
Eastern Kentucky	1.684	0.926	1.767
Campbell	1.732	0.914	1.145
Villanova	1.679	0.911	1.817
Arkansas State	1.773	0.908	1.943

7) Return to the sort dialog and choose “Ascending” and then click “OK.” It now sorts the data in an ascending order, revealing that Lamar has the worst 3-point shot scoring record, with its players averaging 0.325 3-points shots per game. This basically means that only one-third of its players make a 3-point shot each game.

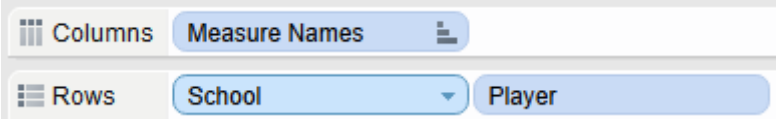
8) To return the table to its original arrangement (alphabetically by school name), click the “School” field under the Rows shelf and select “Clear sort.”



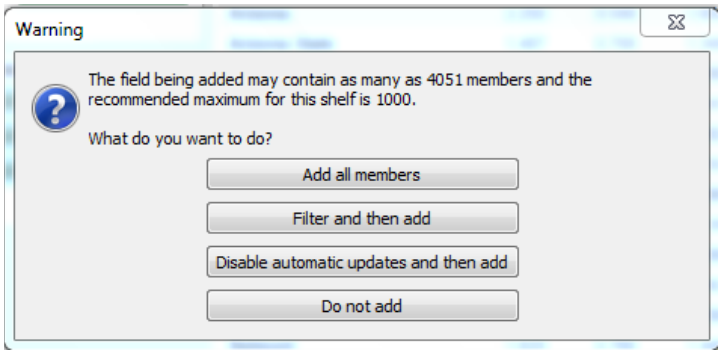
Part 3: Introduce a second dimension to find the best players

We now know which teams' players have the highest average per-game shot records, but we can't differentiate between the best and worst players on those teams. So let's do that.

1) Drag the "Player" dimension to the Rows shelf and place it to the right of "School."



2) It may give you a warning that there are a lot of players that you're adding.

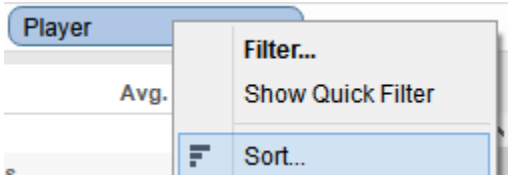


That's ok. Click "Add all members."

3) You'll now see the data organized in two levels, by School first, and then by Player:

School	Player	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
Abilene Christian	Alexsander Milosavljevic	0.895	0.000	0.211
	Austin Cooke	1.065	1.516	1.194
	Chris Blakeley	0.077	0.000	0.000
	Harrison Hawkins	1.875	1.563	3.813
	Jacob Lancaster	2.806	0.032	1.194
	James Pegues	1.742	0.032	1.484
	Julian Edmonson	2.000	1.367	2.967
	LaDarrien Williams	3.188	0.875	3.250
	Michael Grant	1.774	0.000	1.581
	Parker Wentz	1.839	2.548	2.581
	Rafael Farley	0.407	1.222	0.444
	Riley Payne	1.258	0.000	1.032
Air Force	Callum Long	0.286	0.286	0.571
	Chase Kammerer	0.464	0.036	0.393
	Darrius Parker	1.222	0.111	1.056
	DeLovell Earls	1.350	0.200	1.250
	Ethan Michael	0.286	0.143	0.857

4) Right-click on Player and select Sort. As we just did, let's sort this table by average "3p/G" in a descending order. We can see who the best player in each team is in terms of 3-point shots.



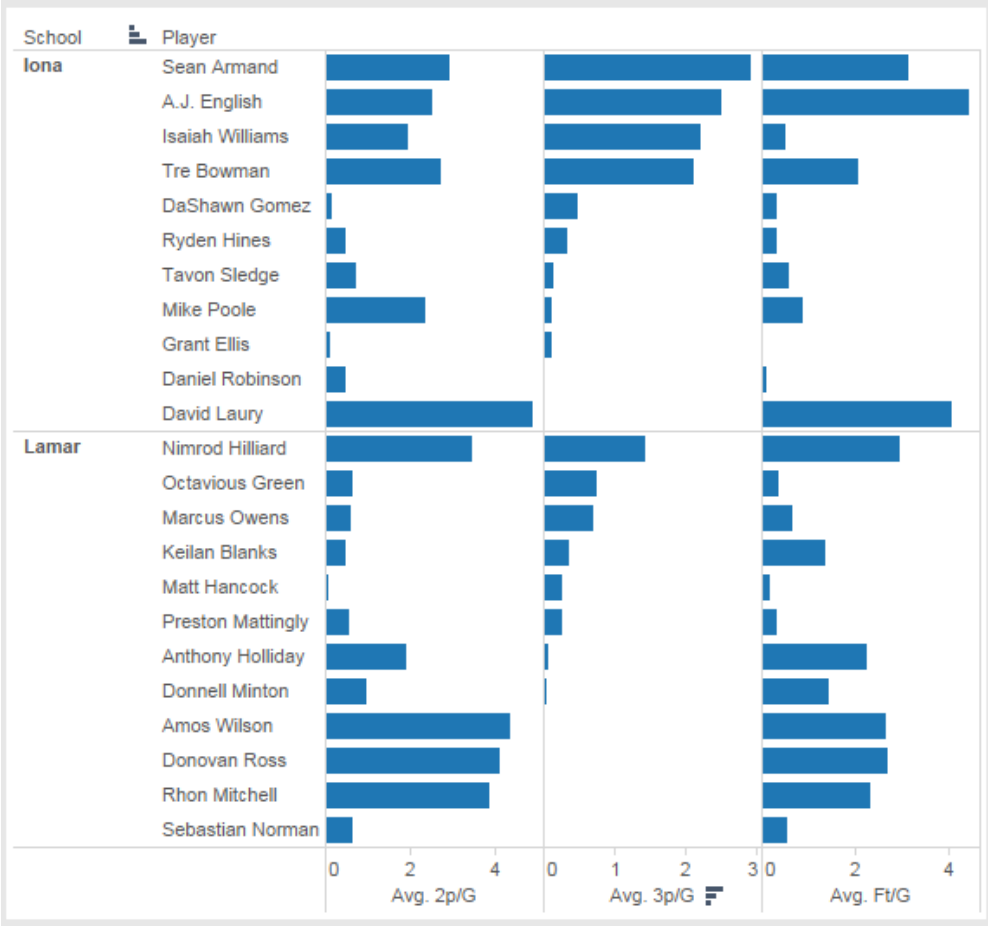
5) Let's take a look at our best and worst 3-point shot schools – Iona and Lamar. Click on "School" in the Rows shelf and select "Filter…"

6) Click "None" to clear all the schools and the select Iona and Lamar. Then click "OK" to apply the filter. You'll see this:

School	Player	Avg. 2p/G	Avg. ..	Avg. Ft/G
Iona	Sean Armand	2.906	2.938	3.156
	A.J. English	2.531	2.531	4.438
	Isaiah Williams	1.963	2.222	0.519
	Tre Bowman	2.710	2.129	2.065
	DaShawn Gomez	0.160	0.480	0.320
	Ryden Hines	0.500	0.350	0.350
	Tavon Sledge	0.741	0.148	0.593
	Mike Poole	2.344	0.125	0.875
	Grant Ellis	0.118	0.118	0.000
	Daniel Robinson	0.500	0.000	0.100
	David Laury	4.875	0.000	4.063
Lamar	Nimrod Hilliard	3.433	1.433	2.967
	Octavious Green	0.640	0.760	0.360
	Marcus Owens	0.630	0.704	0.667
	Keilan Blanks	0.500	0.357	1.357
	Matt Hancock	0.091	0.273	0.182
	Preston Mattingly	0.579	0.263	0.316
	Anthony Holliday	1.929	0.071	2.250
	Donnell Minton	0.957	0.043	1.435
	Amos Wilson	4.367	0.000	2.667
	Donovan Ross	4.100	0.000	2.700
	Rhon Mitchell	3.862	0.000	2.345
	Sebastian Norman	0.654	0.000	0.538

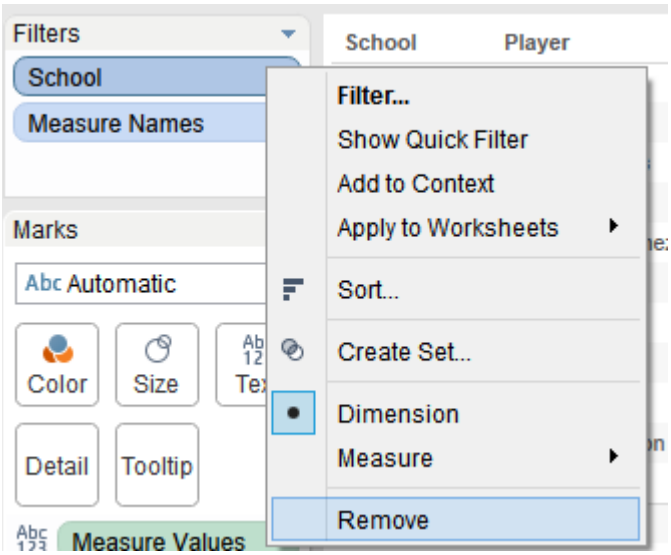
7) You can see that the top 3-point scorer at Iona (Sean Armand) averages twice as many 3-point shots per game as Lamar's top player (Nimrod Hillard). And Iona has two players who haven't scored a 3-point shot this year, while four Lamar players with that distinction.

8) Remember, Tableau still allows you switch easily between visualizations. Try clicking the horizontal bar chart in the "Show Me" area. You'll see this.



9) But we want to stick with text tables, so switch the view back.

10) Now delete "School" from the Filters area (NOT THE ROWS SHELF) to see all the data again.

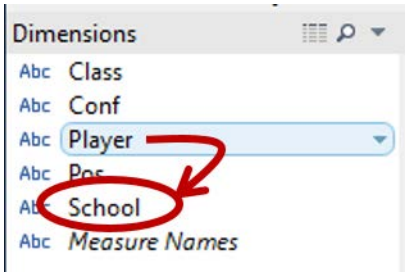


Part 4: Create a hierarchy to enable drilling up (and down)

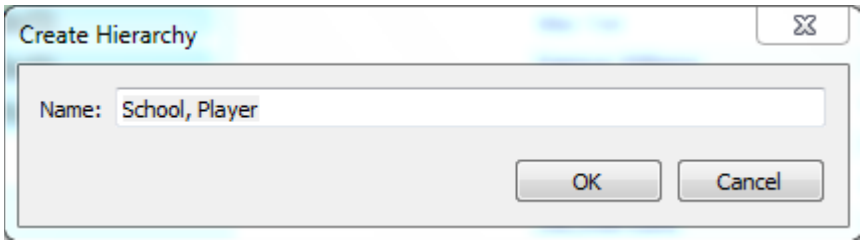
We still have a problem with our table – we lost our school-level averages. We can only see player-level statistics, even though they are organized by school.

We can solve this problem by creating a hierarchy.

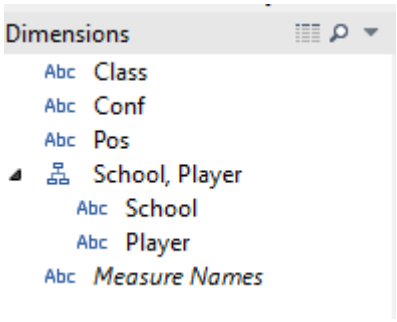
1) Under dimensions, drag "Player" over "School"



2) You'll see a "Create Hierarchy" dialog. Click "OK."

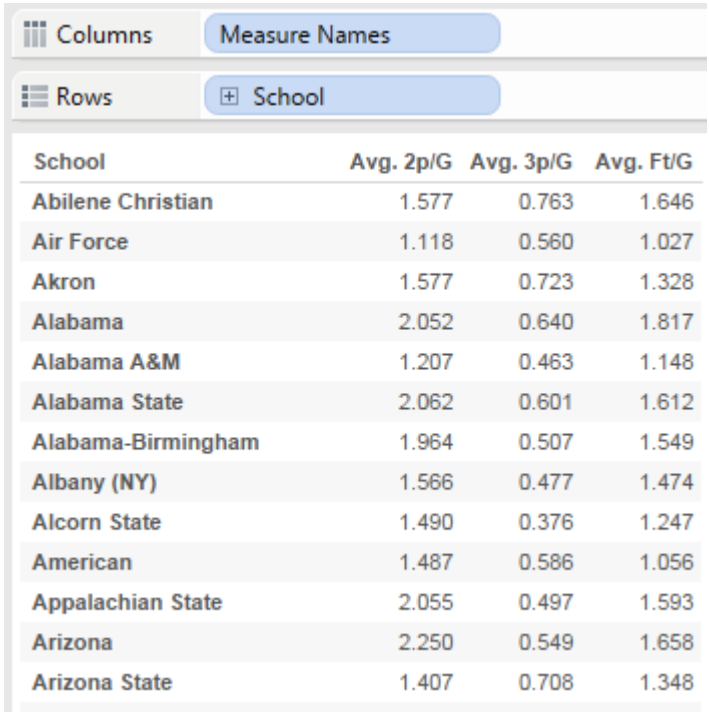


3) It now groups School and Player together into a single structure, where players are part of schools.



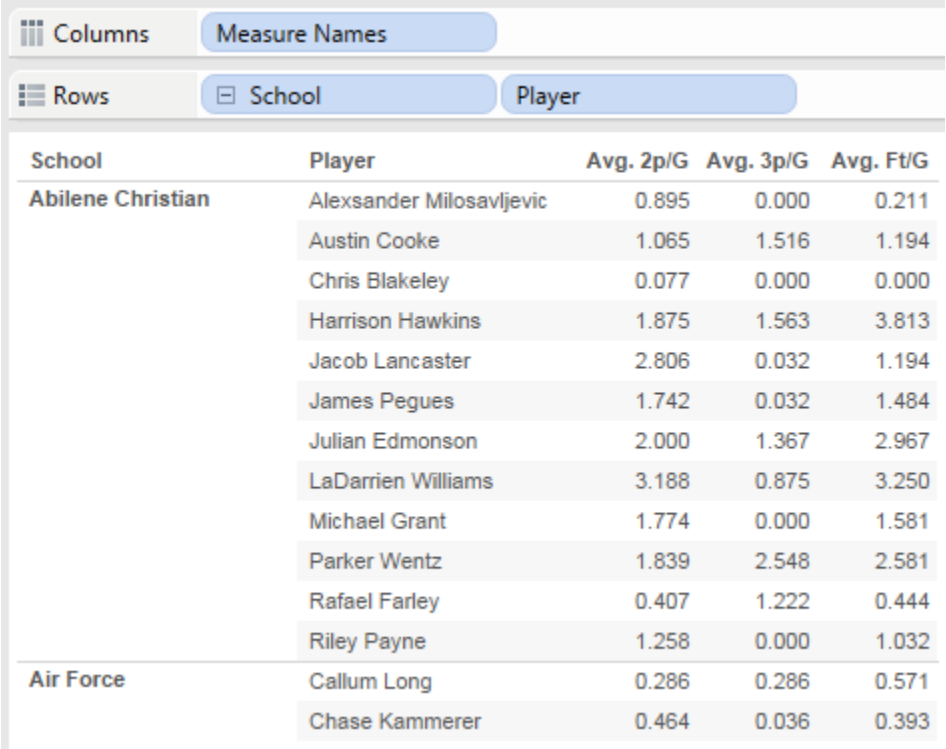
The new hierarchy organizes the data by School first, and then by Player.

4) Now remove “School” and “Player” from the Rows shelf and replace it with the “School, Player” hierarchy. You’ll see this:



School	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
Abilene Christian	1.577	0.763	1.646
Air Force	1.118	0.560	1.027
Akron	1.577	0.723	1.328
Alabama	2.052	0.640	1.817
Alabama A&M	1.207	0.463	1.148
Alabama State	2.062	0.601	1.612
Alabama-Birmingham	1.964	0.507	1.549
Albany (NY)	1.566	0.477	1.474
Alcorn State	1.490	0.376	1.247
American	1.487	0.586	1.056
Appalachian State	2.055	0.497	1.593
Arizona	2.250	0.549	1.658
Arizona State	1.407	0.708	1.348

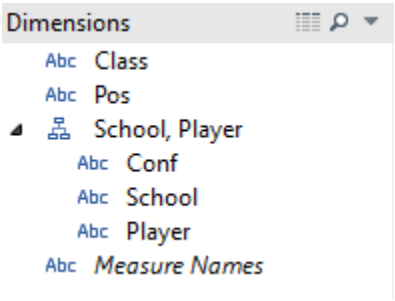
5) Now click the “plus sign” next to “School.” This expands the data to the next level of the hierarchy, allowing you to drill down to see more detail.



School	Player	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
Abilene Christian	Alexsander Milosavljevic	0.895	0.000	0.211
	Austin Cooke	1.065	1.516	1.194
	Chris Blakeley	0.077	0.000	0.000
	Harrison Hawkins	1.875	1.563	3.813
	Jacob Lancaster	2.806	0.032	1.194
	James Pegues	1.742	0.032	1.484
	Julian Edmonson	2.000	1.367	2.967
	LaDarrien Williams	3.188	0.875	3.250
	Michael Grant	1.774	0.000	1.581
	Parker Wentz	1.839	2.548	2.581
	Rafael Farley	0.407	1.222	0.444
	Riley Payne	1.258	0.000	1.032
Air Force	Callum Long	0.286	0.286	0.571
	Chase Kammerer	0.464	0.036	0.393

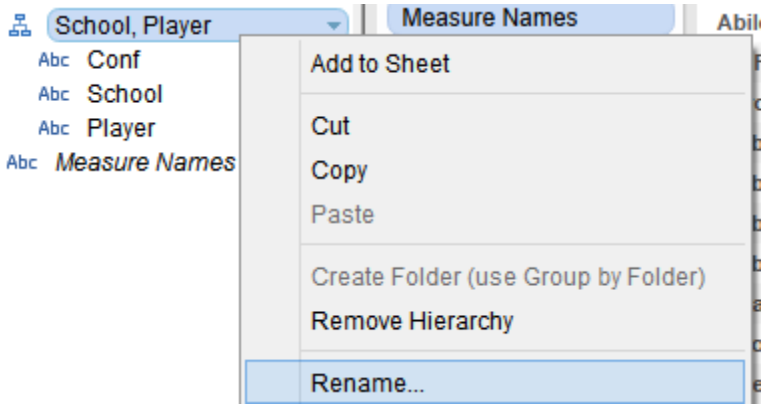
6) Click the “minus sign” next to “School” and it will drill back up to the higher level.

7) Now let’s add a third level to the hierarchy. Drag the “Conf” dimension to the “School, Player” hierarchy. Make sure it is placed above “School.”

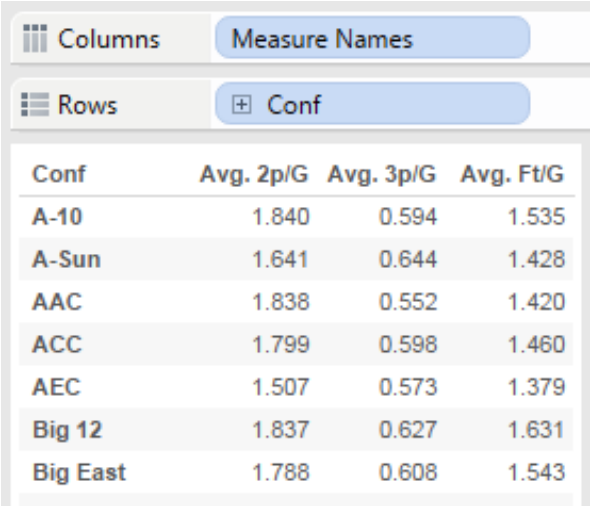


We did this because players are part of schools, and schools are part of conferences.

8) Rename the hierarchy by right-clicking on it and selecting “Rename…” Then change the name to “Conf, School, Player” and click “OK.”



9) Remove “School” and “Player” from the Rows shelf and replace it with the “Conf, School, Player” hierarchy. You’ll now be able to see averages at the conference level.



Conf	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
A-10	1.840	0.594	1.535
A-Sun	1.641	0.644	1.428
AAC	1.838	0.552	1.420
ACC	1.799	0.598	1.460
AEC	1.507	0.573	1.379
Big 12	1.837	0.627	1.631
Big East	1.788	0.608	1.543

or drill down to the player level:

Conf	School	Player	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G
A-10	Dayton	Alex Gavrilovic	1.000	0.071	0.214
		Devin Oliver	3.030	1.091	2.727
		Devon Scott	1.355	0.000	0.935
		Dyshawn Pierre	3.333	0.667	2.394
		Jalen Robinson	1.344	0.156	1.375
		Jordan Sibert	1.788	2.273	2.152
		Kendall Pollard	0.818	0.000	0.394
		Khari Price	0.545	1.333	1.545
		Kyle Davis	0.742	0.032	0.484
		Matt Kavanaugh	2.034	0.000	1.414

Part 5: More with hierarchies

We can still work with the dimensions individually even though they are part of the hierarchy. Let's say we wanted to know who had the best free throw percentage in each conference (NOT in each school).

First, we need to create a calculated field for free throw percentage. The formula is simple.

$$Free\ Throw\ Percentage = \frac{Free\ throws\ made}{Free\ throws\ attempted}$$

1) Select the Analysis menu, and then "Create Calculated Field…"

2) Name the Calculated Field "Free Throw Percentage"

3) For the formula, type.

   [Ft/G]/[Fta/G]

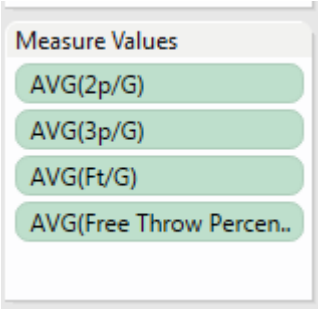
   NOTE: You can double-click the fields to insert them into the Formula text area. Or just type them.

4) Click "OK."

5) Free Throw Percentage will now appear under Measures.

6) Drag “Free Throw Percentage” to the “Measure Values” area (under Marks).

7) Change the calculation from SUM to AVG. You’ll see this:



and this

Conf	School	Player	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G	Avg. Fre..
A-10	Dayton	Alex Gavrilovic	1.000	0.071	0.214	0.600
		Devin Oliver	3.030	1.091	2.727	0.783
		Devon Scott	1.355	0.000	0.935	0.592
		Dyshawn Pierre	3.333	0.667	2.394	0.658
		Jalen Robinson	1.344	0.156	1.375	0.677
		Jordan Sibert	1.788	2.273	2.152	0.710
		Kendall Pollard	0.818	0.000	0.394	0.310
		Khari Price	0.545	1.333	1.545	0.879
		Kyle Davis	0.742	0.032	0.484	0.625
		Matt Kavanaugh	2.034	0.000	1.414	0.672

8) Now remove “School” from the Rows shelf and sort the “Avg. Fre…” column. You’ll see this:

Conf	Player	Avg. 2p/G	Avg. 3p/G	Avg. Ft/G	Avg. ..
A-10	Mike Aaman	1.400	0.000	0.200	1.000
	Seth Berger	0.118	0.000	0.118	1.000
	Papa Ndao	0.267	0.567	0.400	0.923
	Jermaine Myers	0.320	0.200	0.400	0.909
	Austin McBroom	0.906	1.375	1.750	0.903
	Khari Price	0.545	1.333	1.545	0.879

We know that Mike Aaman and Seth Berger led the A-10 conference in free throw success (both made 100% of their attempts). This doesn’t mean that they had the highest number of successful free throws, just that their attempts were the most successful.

Part 6: Try it yourself!

Answer the questions below. You should remove measures and dimensions from the Measure Values area and the Columns and Rows shelves so that your workspace isn't too cluttered.

Note: This data set indicates players that play multiple positions (i.e., G-F is Guard & Forward). For the purposes of this analysis, you can ignore those.

1) Which team's players averaged the most points per game?

2) What was the average points per game for a Temple player?

3) Do senior (SR) forwards (F) have a higher free throw percentage than freshmen (FR) centers (C)?
   HINT: Construct a hierarchy using Class and Pos.

4) Do players appear to get better at free throws (a higher percentage of their attempts are successful) as they get older? Provide evidence.

5) Are forwards (F), in general, better at free throws than centers (C)? Provide evidence.

Analyze the 2013-2014 Temple team:

HINT: (apply a filter so you only see Temple data)

6) Which class' players (FR, SO, JR, SR) averaged the most 2-point shot **attempts** per game (not 2-point shot scored)? How many did they average?

7) Which class' players averaged the least 2-point shot attempts per game? How many did they average?

8) Analyze the average minutes per game that players spend by class and determine whether that explains what you found in Q6 and Q7.