

Comp 4433: Assignment 3

When viewing scatter plots, we are often primarily concerned with the general pattern of the relationship between two variables. In other words, we are interested in the relationship as a whole. The scatter plot can also be useful for case-level analysis, where we are interested in individual points as opposed to the larger trend.

The quadrant plot, a specific application of a scatter plot, is an example of a case-focused scatter. This plot uses a pair of intersecting horizontal and vertical lines to decompose the plotting space into four quadrants. A Cartesian plane, of course, naturally has four quadrants established by the x and y axes. We can use the natural quadrants established by the x and y axes to form our quadrants, or we can establish our own horizontal and vertical thresholds based on meaningful values. When data are converted to standard scores (z-scores) then using the x and y axes to form the quadrants often makes sense, as the positioning of a given point will indicate whether that observation is above or below the mean for each of the two values being scattered.

Using the provided dataset, NBA.csv, you will build two new metrics, one to represent offensive performance and the other to represent defensive performance. These scores may be simple weighted linear combinations of certain variables that you deem important or they may be more complex, such as the first principal component from a PCA.

After establishing your new metrics, use them as the basis for a quadrant plot in Plotly. The results should allow the user to see relative player performance based on your performance metrics.

Be sure to include:

- Visible quadrant boundaries
- Some visual mechanism to distinguish between the points falling into each quadrant.
- Labels for each quadrant
- Tooltip information indicating the player represented by each point

Provide a short explanation of the plot via a markdown cell in your notebook.

You may use any Plotly functionality to accomplish this task and are not limited to content discussed in class. Note that the dataset contains multiple records for some players as a result of mid-season trades. Feel free to deal with this duplication in any way you prefer.

Descriptions of the features are available at the following link:

https://www.basketball-reference.com/leagues/NBA_2024_per_game.html