

Athletic Data Analysis

Ishani Desai

Mentor: Loretta Sue Auvil

Department of Statistics and Computer Science, College of Liberal Arts and Sciences, University of Illinois at Urbana-Champaign

INTRODUCTION

Our initial research involved finding tendencies in opposing teams’ play-calling, such as play outcomes differing based on certain players being lined up in certain positions. These tendencies are found by generating frequent patterns and association rules from play-level game data. Visualizing these findings is an important next step; a good visualization can make it easier to understand which rules are strong, useful, or relevant, as well as any patterns between the rules, which can help coaches for game planning decisions.

AIM

Our primary goal this semester was to visualize patterns and tendencies we’ve found in football play data. Our aim was to create a visualization that the reader can create a clear narrative from. We also wanted to ensure that reader can easily pick out a rule based on its features or strength, and easily understand it. The data from which the itemsets and rules were generated was from Pro Football Focus, a company that generates advanced stats on the play level, as well as from Illini Football. The rules were generated from this data through an algorithm that was initially created by my mentor, which I worked on this past summer and fall.

METHOD

The figure above is a visualization of similar data. Each row on the left-hand side represents a pattern, and the columns each represent a feature. We wanted to create a visualization with similar axes, but with more interactive elements.

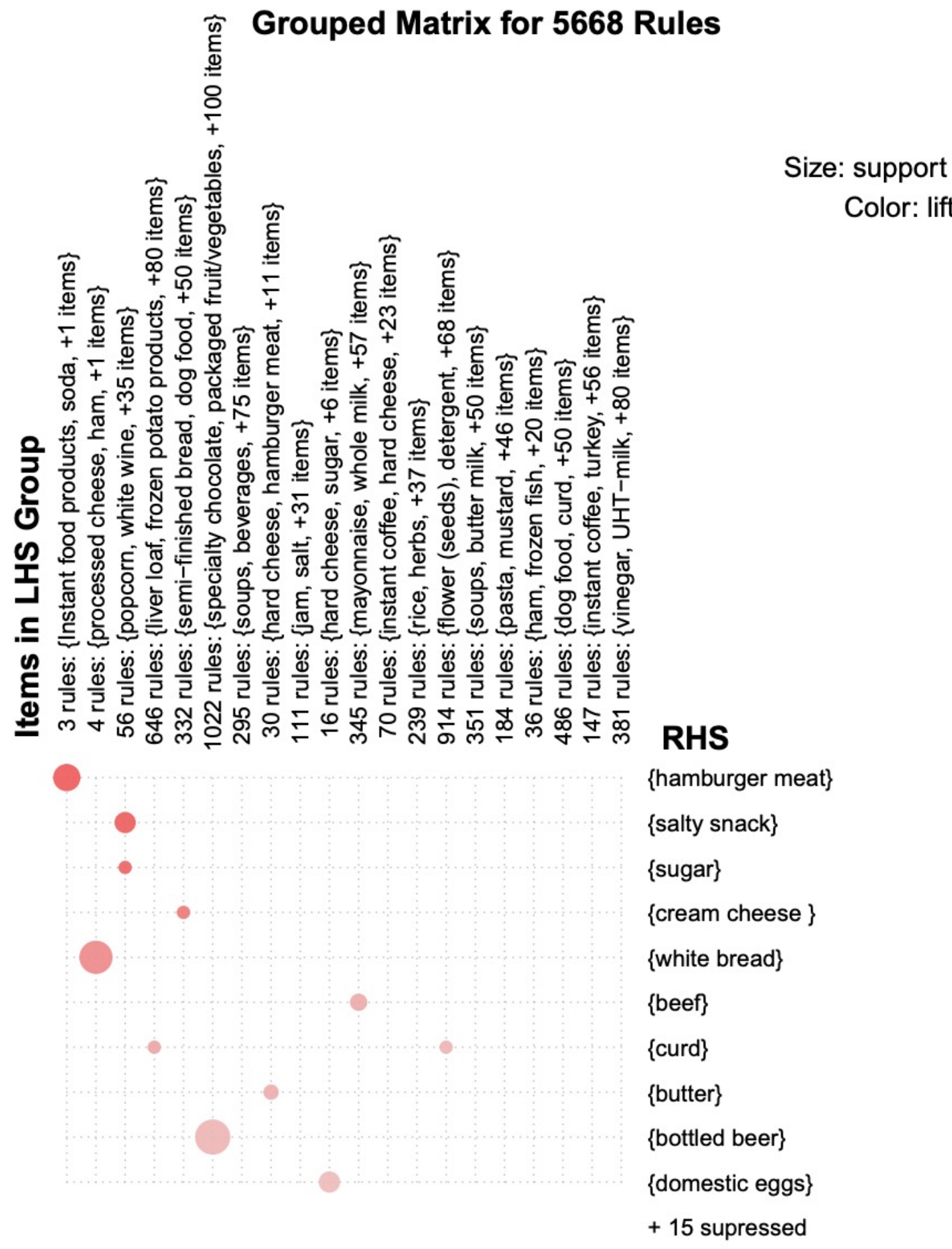
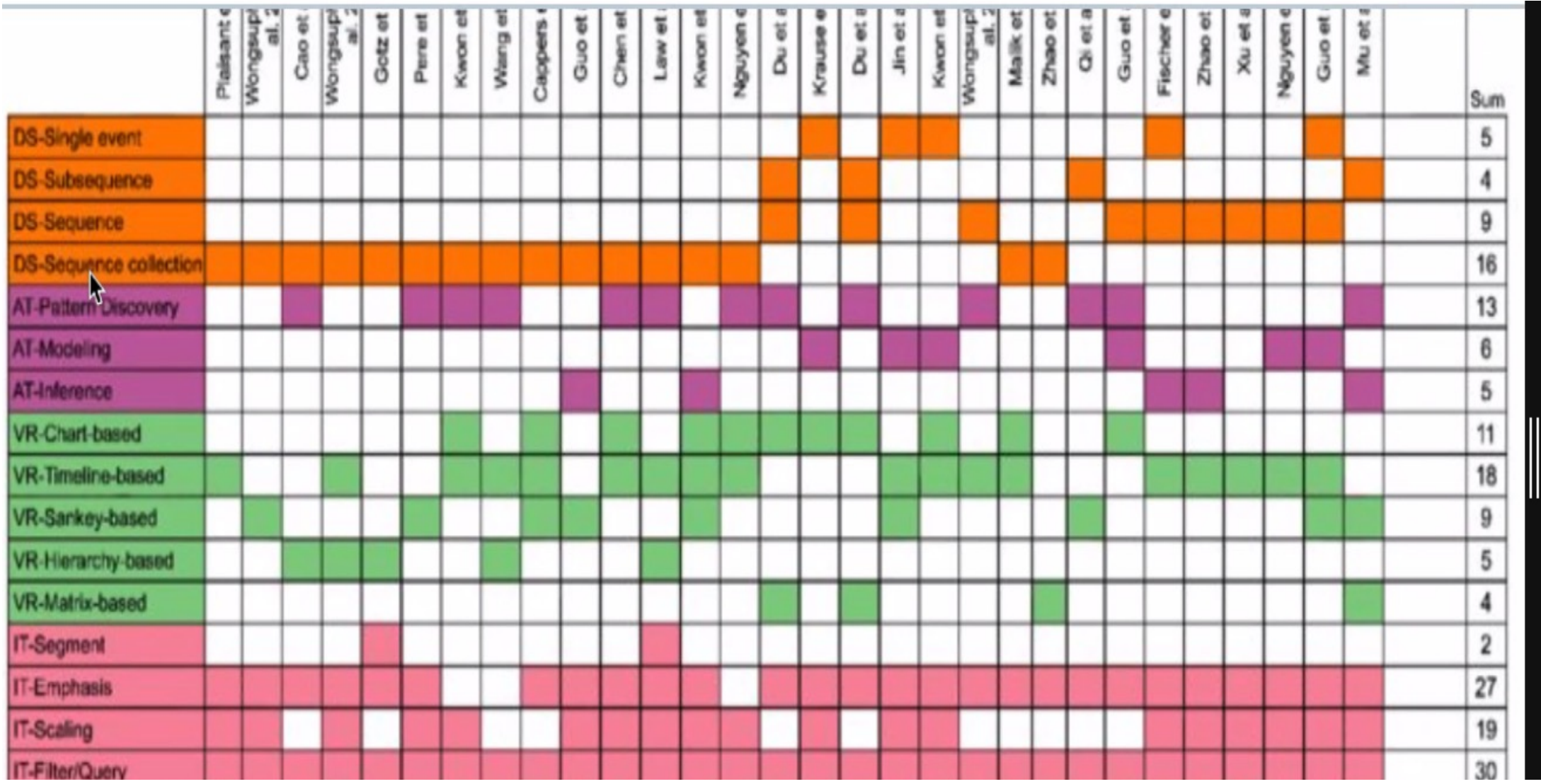


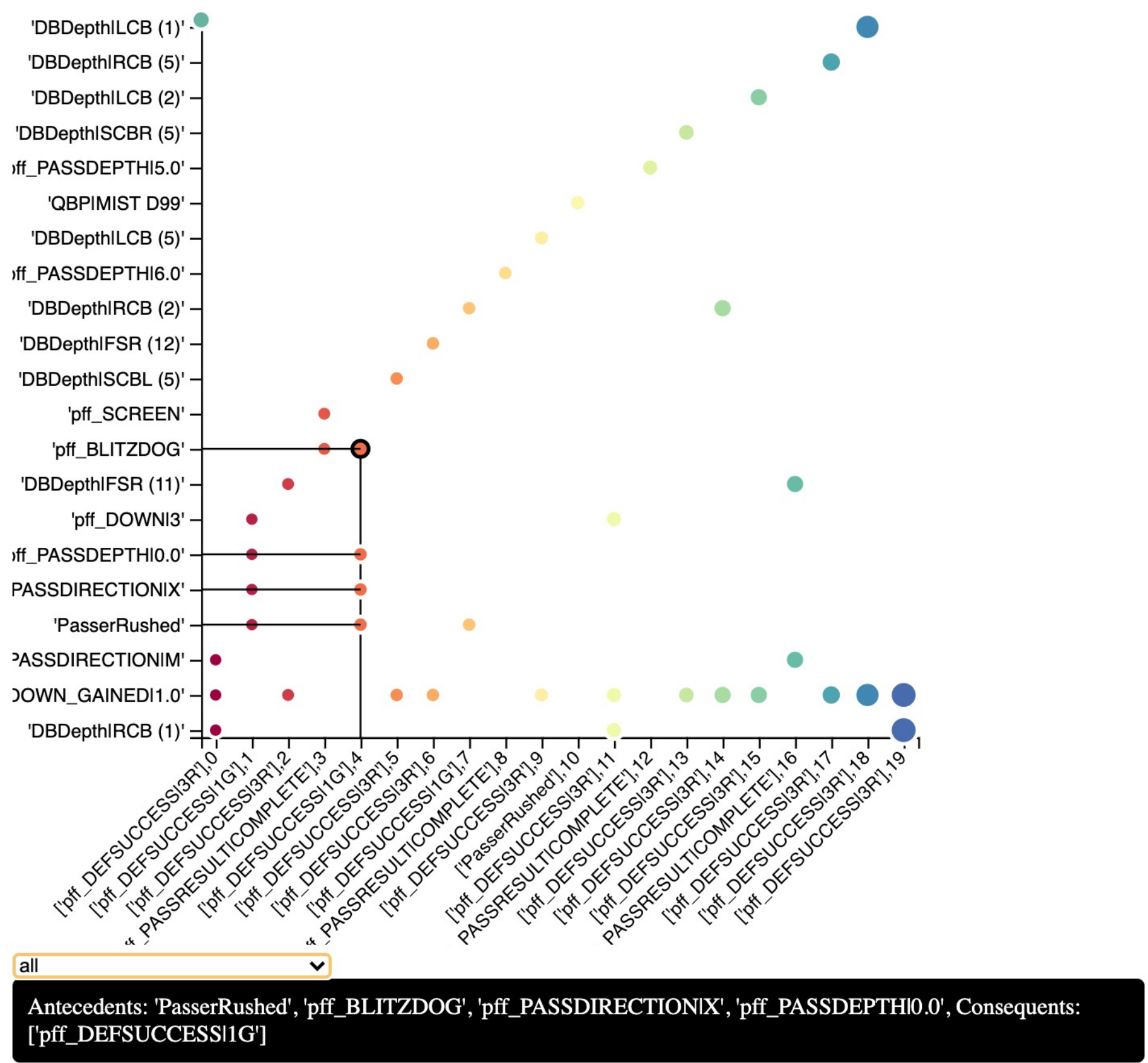
Figure 7: Grouped matrix-based visualization.

The above visualization is similar to the first, with rules grouped as each axis point on the x-axis, and features on the y-axis. The size of the bubbles represents the rule strength, which I wanted to incorporate into this visualization. However, we wanted to avoid grouping features, such as the above figure does in the x-axis, for readability.



RESULTS

We created an interactive visualization with points representing the features in a rule. Each column represents an individual rule where the x-axis includes the rule number and its consequent, while the y-axis consists of all the antecedents. When the mouse is hovered over a point, the visualization indicates the other antecedents in the rule and which consequent they belong to, and the full info is included below. The color indicates which points belong to the same rule, and the size indicates the strength of the rule. There is also a dropdown bar to filter based on features.



RECOMMENDATIONS

There is a lot of future work that can be done to improve this visualization. A dual x-axis can be implemented to group together rules with the same consequent, instead of both the rule number and consequent being on each tick mark. The filtering options can also be expanded to include rule strength (confidence and support). We would also like to fix the coloring system to better distinguish which points belong to the same rule and make smaller aesthetic improvements to the axes and plot.

CONCLUSIONS

Our pattern analysis yields important insights about play-calling, and a visualization can help to communicate these insights to coaching staffs and increase practical use of our work. This visualization can help the user pick out which rules are useful to them depending on what kinds of features they are interested in examining.

ACKNOWLEDGEMENTS

I’d like to acknowledge the National Center for Supercomputing Applications SPIN program for supporting this project. I’d also like to thank my mentors Loretta Sue Auvil and Kingsley Osei-Asibey for guiding me through my research.