Web-based Visual Analytics for Player Performance in Fantasy Football

Jonny Moonyham[1], Austin Corotan[2], Scott Waldron[3]

***Abstract*— The methods and visualizations described in this paper aim to provide increased insight into the process of adding or dropping a fantasy football athlete during the regular season. The goal is to improve user experience for new users while remaining novel for those veteran to the game of fantasy football. The proposed visualizations combine individual player performance with team composition in a new format to better understand the trends of athletes. The visualizations were evaluated with the help of experienced fantasy football participants. It is hoped to extend these visualizations in an open source environment to remain free to the public.**

I. Introduction

Fantasy sports started as a game between friends but soon became a multi-billion dollar business influencing and evolving sports. The thrill of fantasy sports is quick to pull in new players and keep the existing player-base content with new content every season. Participants create a virtual team from real players and compete against other participants in the same league. The team that ends the week with more points than their opponent is the victor. Players can track their virtual team's points and participation through multiple websites and apps gaining insight to which athletes are performing well and which ones should be replaced. There are approximately 59 million participants playing fantasy sports in the United States and Canada with roughly 40 million participating in fantasy football. With an ever-growing data set involving fantasy sport statistics, the focus of this paper is centralized around fantasy football and the performance of NFL athletes.

Currently, fantasy football statistics and metrics reside primarily in tables and charts requiring cognitive and experienced interpretation to extract meaningful evaluation. Ultimately, participants are looking for key markers in players that indicate a favorable addition to their team. Multiple websites and visual tools exist to help users quickly identify key players, but these sites are fragmented and frequently only offer their services behind a paid subscription. This paper focuses on two types of users: the experienced fantasy football participants understanding the inner workings of fantasy football, and those who are new to the game and are looking to learn. Using publicly available JavaScript libraries and a local web hosting service, the visualizations provided aim to answer some very simply questions. As a broad, overarching subject, do the visualizations improve user experience interaction with fantasy football data? And more pointedly, do the visualizations provide meaningful insight into the process of adding or dropping a player based on prior performance and weekly match-up?

Projected Points are an indicator of a player's anticipated performance in the coming week and have a substantial impact on player opinion. Projections can either be computed using proprietary algorithms or issued by professionals. Currently the projections only exist on major platforms offering the most recent projected data point. If players are wanting to extract a meaningful correlation between projected and accrued points, they are required to dig through past tables and construct their own relationship. One solution this paper offers is to provide users with the ability to quickly identify a correlation between projections and accruals in a quick and concise format for a given season.

Players can accumulate points in a variety of ways in fantasy football. They can catch, run, or kick the ball or possibly score a touchdown. Quarterbacks throw for passing yards, running backs run for rushing yards, and kickers kick for field goals. With a visualization tool to break down where a football player accrued their points, participants could effortlessly determine where a player receives their majority. If there exists some running back that establishes high points on a weekly basis, the visualization would quickly identify if the running back is touchdown dependent or if they indeed rush for many yards. If a player is touchdown dependent, it could lead to further insight when the rushing team is up against an experienced defense not allowing many points scored.

It is also important to note how a team works as a composition. Not only is it important to look at a players performance, but the performance of a team as a whole. If a participant of fantasy football is interested in adding a player to their roster, they not only have to consider individual performance but team performance as well. With an implementation of a team composition visualization, players would be able to efficiently discern where a team excels and who the key players are within the excellency. By allowing the visualization to be interactive, those seeking information on a weekly basis can scroll week by week to discover a more in depth team composition break down.

Lastly, understanding player depth and injury status on a team is a key factor in the decision process of adding a player to a roster. Players considered first-string will undoubtedly be on the field more often then substitute players and even more so than possibly injured players. These players, through more exposure on the field, often enjoy a higher fantasy score. Introducing a visualization that can help fantasy football

[1]J. Mooneyham is a graduate student at Western Washington University pursuing his Master of Science in Computer Science.

[2]A. Corotan is a graduate student at Western Washington University pursuing his Master of Science in Computer Science.

[3]S. Waldron is a graduate student at Western Washington University pursuing his Master of Science in Computer Science.

participants quickly identify player depth and injury status on a weekly basis can lead to an increased ability to make correct player acquisitions when it comes to game day.

The aim of this paper, with all of the visualizations presented and discussed in greater detail later, is to make fantasy football more presentable and accessible to new users while providing novelty and increased judgment to veteran players when adding or dropping players from a fantasy football lineup. With an open source platform and no subscription required, these visualizations are available to a user with a web browser.

II. Methods

A. Tools

For our project, we used Python 3 to support data collection and coalescing. The Beautiful Soup Python library was used to scrape data from websites. Python algorithms were used to shape collected data into appropriate JSON and CSV formats. All coding implementation can be found here: https://github.com/mooneyj3/cs597data.

For the visualizations we used D3 (v5) and Tableau Public (v10.5). D3 provided a rich environment to develop customizable interactive visualizations. The use of Tableau Public enabled quick prototyping of common idioms that are embeddable within a webpage. Visualizations were embedded in a web platform driven by VueJS (v2) components. The web platform was structured with Vue Router for navigation and Vuetify for elemental design layouts. With the support of public examples and D3 library documentation, we were able to generate visualizations and tailor them to our tasks. The full web implementation can be found here: https://github.com/mooneyj3/cs597project.

B. Data Collection and Processing

We selected the first 8 weeks of the 2018 NFL Regular season as a restricted time range for data collection. This served as a reasonable window to collect data and make comparisons in parallel with a season currently in progress. Data had to be collected from multiple sources to support visualization tasks.

1) ESPN Player Predictions: In the repository for data collection, *playerDataScraper.py* was used to scrape the data from the tables of ESPN using Python's BeautifulSoup library. BeautifulSoup is a library to abstract away the naming conventions of *HTML* and *XPath* leaving the user with a quick and informative set of library calls to extract information from a python http request. The information presented itself in numerous tables and links that required a recursive solution to examine all player statistics on a weekly basis. Once projection data was extracted for the players, the program had to navigate to another ESPN domain to gather the accrued points to pair the two sets together in a library. *HTML* Scraping was intuitive and overall helpful as it allowed the format customization of the delivered JSON file.

2) NFL.com JSON Game Summaries: From the cs597data repository mentioned above, three main Python scripts were used to collect and process the data. *play_by_play_fetcher.py* is a customized script that fetches raw JSON data for each game in the specified time window. Approximately 110 files were collected (12.9 MB in total). Each file represents a game between two teams contains a detailed breakdown of player statistics for each player and a summary of each play in the game. These JSON files contained too much hierarchical depth and needed to be processed for simplification.

The first script, *pbp_to_team_summary.py* iterates through the files to process the data into team summaries. This generates files per team, that has a detailed breakdown of all the player statistics per team. The resulting files were effective for data exploration and initial prototyping. A follow-up script, *modify_team_stats.py*, was developed to generate CSV files that could be easily used with D3 and Tableau Public.

3) Depth and Injury Data from sportsradar.com: Depth and Injury data were retrieved from the sportsradar.com API in JSON format. Users can specify the season, whether or not they want regular or post season, and the week. The JSON files are formatted by team, position then injuries/depth at that position. Details of the data retrieval process using the sportsradar API are shown in the *depth_injury_api.py* script found in the cs597data repository.

III. Visualizations

Here we describe the specific visualizations we implemented. Interactive visualizations are viewable at https://dataviz.moonham.com/players/1234.

A. Projection Evaluation

A key component to understanding player performance in fantasy football is to understand if players outperform their projected values or if they consistently under perform. Projected points are often linked to proprietary algorithms or professional opinions and can differ tremendously on a weekly basis. To combat the convolution of projected data against accrued data, the projection evaluation visualization (see **Fig 1.**) establishes a format to quickly identify a player's performance with regards to their projected value.

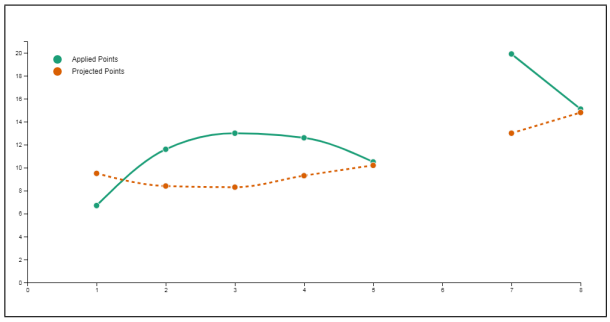

Fig. 1. Interactive line graph showing accrued points with projected points. The axes are points vs week

By using a simple line chart idiom, users can quickly identify a player's performance on a weekly basis. With current week of the season shown along the xaxis and fantasy points shown on the yaxis, the graph can not only visualize accrued points but projected points as well. Hue and shape, with the use of a legend, are used to show distinction between projected points and accrued points. With a bye week for the player on week six, the chart removes the data point and continues the following week. This graph also includes interactivity with a hover feature. When users hover over data points, they get the concise value for that data point instead of having to estimate based on the yaxis.

This visualization was created using the D3 library in Javascript. By reading and understanding multiple line graph solutions from *bl.ocks.org* such as *D3 v5 Line Chart* and *Reusable Responsive Multiline Chart*, this graph was made from the tailored JSON file described earlier.

This graph doesn't offer player to player comparisons currently but offers quick understanding of the relationship of one player's performance with their projections. As we'll see later, future work includes the player to player comparison with the option to add and remove numerous players at once and the ability to toggle projections on and off.

B. Past Performance Fantasy Point Breakdown

To support users with analyzing the past performance of a player we created the Past Performance Fantasy Point Breakdown visualization (see **Fig. 2**). This helps a user see which categories contribute to a players overall fantasy points by week. This is a helpful companion visualization to the projection evaluation visualization allowing a user to explore the source of player points.

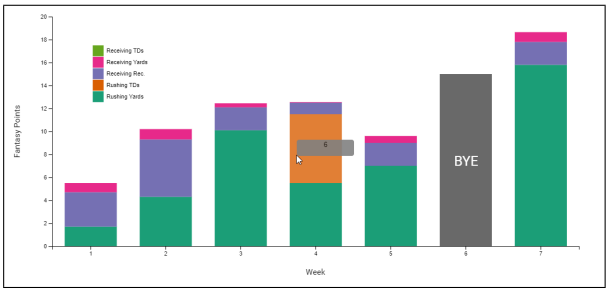

Fig. 2. Interactive stacked bar chart showing oints with projected points. The axes are points vs week

For this visualization, we selected the stacked bar chart idiom. Week of the season is displayed in the xaxis and the Fantasy Points is represented in the yaxis. The height of each column along the xaxis shows the total fantasy points. Hue is used to encode the different categories that contribute to total fantasy points, and the relative height of each category allows for comparison. A gray block was selected to represent the BYE week for the player. Interactivity allows a user to hover over different parts of the graph to see actual fantasy point values.

The visualization was implemented using D3. We followed code provided by Mike Bostock to create the initial layout for this visualization (see https://beta.observablehq.com/@mbostock/d3-stacked-area-chart). It was further enhanced to support interactivity and other features unique to our project.

A user of this visualization will be able to compare point values for different categories from week to week for the selected player. Additionally, this enables a user to explore trends with the selected player at a categorical level. While this doesnt support cross-player comparisons, it helps users decide if a player is worth adding to their team.

C. Team Treemap for Comparisons

With our team treemap visualization (Fig. 3), users are able to analyze the composition of fantasy points within the selected players team. This will help a user understand the point contribution of the selected player to their team, the point contribution of other players, and the point contribution of the team as a whole. A user can view this for the whole season or select a specific week.

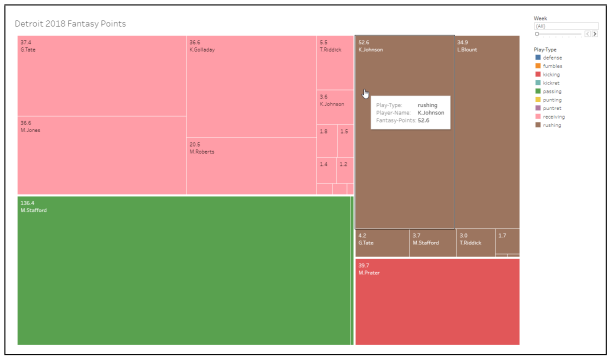

Fig. 3. Interactive stacked bar chart showing oints with projected points. The axes are points vs week

For this visualization, we selected the treemap idiom using Tableau Public. This visualization was not originally intended to be a deliverable. While exploring data in Tableau Public, we found the treemap presented an effective visualization that spans a broad category of details. Color is used to differentiate play categories, for example receiving points versus rushing points. Size is used to denote the overall fantasy point contribution of a player within a space. Ordering is also an important aspect of this idiom, so we see the highest scoring players positioned towards the top-left most portions of each subgraph.

Interactivity is an important part of this visualization given the complexity and overall quantity of the data. Users can over over rectangles to get details. This is especially helpful when players with smaller point contributions are contained in smaller rectangles. Clicking on a rectangle brings a player into focus, reducing the luminance on all other players. Filtering can also be applied. At this time, it only allows filtering by week and by play-type.

By using the treemap idiom, users are able to explore the selected players team as a whole. This is a novel use of visualization in fantasy football, that has not been applied from what we have seen. By exploring team composition, a

user can develop an understanding of player strengths and offensive tactics. This visualization can also be compared with the team depth chart visualization to compare how injuries affect a team and overall points.

D. Team Depth Chart

Our depth and injury visualization (Fig. 4) allows users to analyze the status of two teams by player availability. Specifically, the visualization shows the depth and injury information for players on the team of the desired player and players on the opposing team. The primary goal of this visualization is to provide users with insight on how a desired player might perform based on their availability in the lineup and the availability of players that might influence the targeted player position.

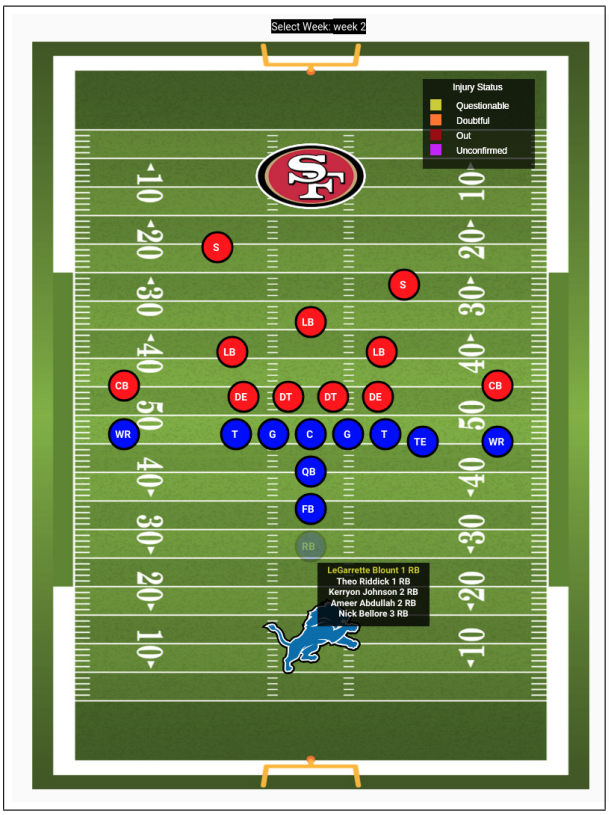

Fig. 4. Interactive depth and injury chart showing tooltip mouse-over. Users can select their desired week.

The visualization was implemented using D3 and uses a custom visualization idiom. The primary marks are points and the primary channels are position and color. Player positions are contextualized by overlaying them on a diagram of a football field. Red marks indicated defensive team positions and blue marks indicate the offensive team positions. When a user hovers over a position a tooltip will list players according to depth, with the uppermost player being first on the depth chart. Players names will be highlighted a certain color based on their injury status.

The rationale for our design techniques regarding this visualization are mainly to contextualize the overall team status as well as the status of the desired player. Providing this information will give users the ability to determine how other players might affect the performance of the desired player in addition their past statistics. Player status information is typically presented in a tabular format and users have little to no idea of how that affects the team as a whole. This visualization provides a novel approach to understanding and exploring this information further.

IV. Findings

The three main questions that our paper aimed to answer were as follows:

- *Which player(s) at a given position will score the most fantasy points this week?*
- *What player(s) on my team will perform the worst this week?*
- *Which statistics best determine player performance on a weekly basis?*

Our visualizations were evaluated through case studies with fantasy football players. We provided players with a specific scenario and associated tasks. Using our visualizations, players had to complete the designated tasks and provide feedback on their process.

The scenario provided was focused on a specific player, *Kerryon Johnson*. Users had to determine if he was a worthwhile pickup for their teams. We provided a set of questions in the form of a survey to detail the experience that the users had with our visualizations and the task. The questions were as follows:

- *How well do the visualizations help you make an informed decision about adding Kerryon Johnson to your team?*
- *How helpful are the visualizations?*
- *How informative are the visualizations?*
- *What parts of the visualizations were the least helpful/informative?*
- *What parts were the most helpful?*
- *Which components, if any, felt unclear?*
- *What changes can be made to improve existing visualizations?*
- *How likely are you to use these visualizations to determine player performance?*
- *How likely are you to recommend these visualization tools to a friend new to fantasy football?*
- *How likely are you to recommend these visualization tools to a friend very familiar with fantasy football?*
- *Are there any data visualizations missing that would help you make a more informed decision?*

Users were given a 1-10 scoring system where one is the lowest score, or ”least likely” action and ten is the highest score or ”most likely” action. Questions were also open-ended to allow users to explain in detail their feedback of the visualizations. We received a total of 14 responses for users across varying fantasy football leagues.

A. Evaluation Results

Nearly 72% of users rated our visualizations eight and above when answering the question: *How well do the*

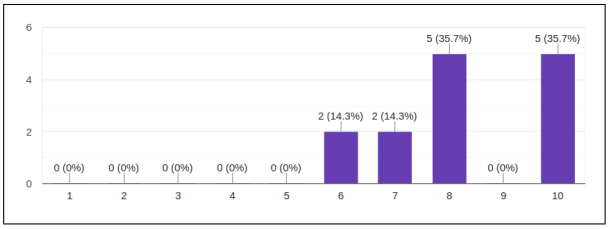

Fig. 5. Feedback: How well do the visualizations help you make an informed decision about adding Kerryon Johnson to your team?

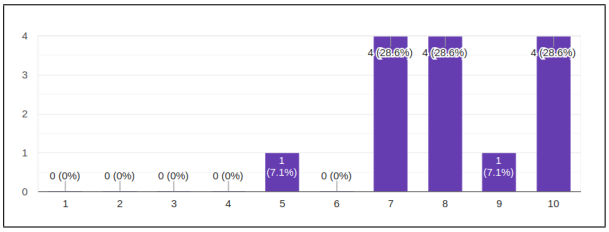

Fig. 6. Feedback: How helpful are the visualizations?

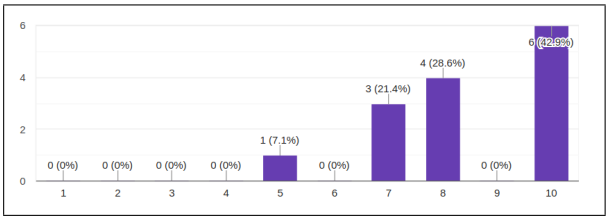

Fig. 7. Feedback: How informative are the visualizations?

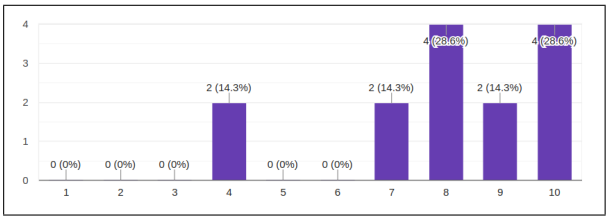

Fig. 8. Feedback: How likely are you to use these visualizations to determine player performance?

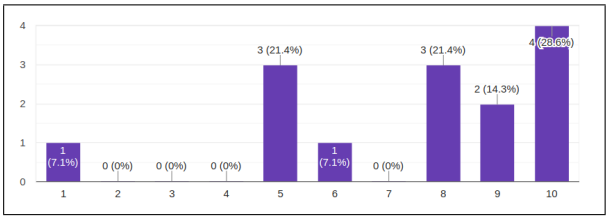

Fig. 9. Feedback: How likely are you to recommend these visualization tools to a friend new to fantasy football?

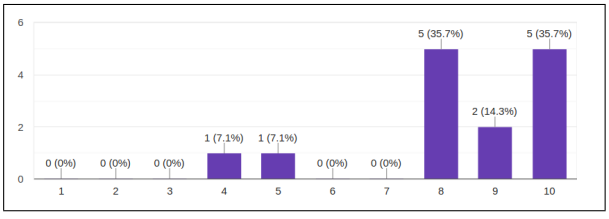

Fig. 10. Feedback: How likely are you to recommend these visualization tools to a friend very familiar with fantasy football?

visualizations help you make an informed decision about adding Kerryon Johnson to your team? (Fig. 5). Almost 65% of users rated our visualizations eight or higher based on how helpful they thought they were (Fig. 6). About 72% of users gave our visualizations an eight or higher based on how informative they thought they were (Fig. 7). 72% of users said they were very likely going to use these visualizations to determine player performance, rating above an eight (Fig. 8). Finally, almost 65% of users believe they will likely recommend our visualizations to new players (Fig. 9) and 86% of users believe they will likely recommend our visualizations to experienced players (Fig. 10).

In addition to these metrics, we also received detailed feedback regarding the strengths and limitations of our visualizations. Some of the most helpful responses are as follows:

Question: Which components were the least helpful/informative?

- *I have doubts the depth chart visualization would stay up to date. Too much changes too fast and I don't know what the data source is.*
- *The last two aren't that helpful as the team score doesn't mean much for an individual player a lot of the time, and the depth chart can change over time, and this one looks like it from the beginning of the year.*

Question: Which components were the most helpful/informative?

- *I really appreciated the team fantasy points graph, that can help in a lot of ways. I also liked the first graph with the applied points and projected points. It's a really simple, quick way to evaluate a player and makes a decision much easier.*
- *I really like the bar graph because I don't follow football, and make all of my decisions off of past and predicted points.*

Question: What changes can be made to improve the existing visualizations?

- *I don't know if this is possible, but if I could compare two players side by side that would be really helpful.*
- *It would be helpful if the second one had the horizontal gridlines visible in the background so that it was easier to identify a single unit with the naked eye.*

Question: Are there any data visualizations missing that would help you make a more informed decision?

- *A chart that showed the RB points allowed to each opposing RB per week by the upcoming opponent. - How has this Defense done against the RB each week so far, and who were those other RBs each week?*
- *A chart very similar to the first one, but that also included the other players of the same position on the same team each week. - How does this player compare against the other players he's competing for field time?*

Based on the initial feedback provided by our 14 users, majority felt that they could sufficiently complete the pro-

vided task. Users thought that the *Player Projection* line graph was clear and easy to understand. They also found the *Fantasy Point Breakdown* stacked bar chart useful in getting a quickly glance at the breakdown of past player statistics. The main visualization users had trouble understanding was the *Depth and Injury* visualization. This is not a surprise because this is a custom visualization and one that requires background experience in both football and fantasy football. The main issue was that users were unaware that the visualization was interactive at first. The main suggestions for improvement of this visualization is to add a snippet notifying users how the visualization works. Finally, the majority of the users felt that they could recommend our visualizations to both new and experienced players.

V. Discussion

Data collection is one of the most significant challenges in this project and going forward. A majority of quality NFL data is behind a paywall through an NFL partner. Access to the data is very restrictive, and costs tens of thousands of dollar per year. It is for this reason, we had to find other means such as scraping, free data sources, and third-party APIs. With access to NFLs API, we would be able to alleviate some of the tedium.

Another key challenge is coalescing the data. Pulling data from different sources prevented alignment. An example is ESPN player identification numbers differing from those attached to NFL data. The issue is exacerbated when NFL presents their player names with first initial followed by last name, while ESPN shows full name. Thus, coalescing requires additional effort to bring beyond what we expected.

We also faced challenges in developing the visualizations with D3 trying to get chart sizing to be managed relative to the view. We have not addressed this yet, however, the visualizations in their current form cannot be dynamically sized, and do not scale well to a smaller form. This will be an important aspect in our future work, as we need to be able to display more visualizations on a single page to better communicate all the characteristics of a selected player.

A limitation of our data is the lack of interactivity between players. This is a common challenge with football data, as different players and different formations result in a variety of plays. It is possible to obtain this data through (1) the pay-per-use NFL API discussed above, or (2) through natural language processing on play-by-play data that is free from the NFL API.

A limitation in our visualizations is the use of cohesive colors and cross-visualization interactivity. Since the visualizations were developed independently with different tools and templates, color was not managed consistently. Additionally, cross visualization interactivity could be increased given that week is a data feature in each visualization.

Another limitation in our visualizations is the lack of comparisons to other players. Often a fantasy players need a relative comparison for value. Adding and dropping a player typically represents a trade-off. This is a shortcoming of our visualizations, that was also a source of repeated feedback from individuals from our test group. This limitation was primarily an issue with data coalescing and architecture. Data for the prototype visualizations represented a static export of data from our snapshot time frame with limited coalescing. An ideal system would have a robust backend architecture with database and restful services. Altogether, this is a limitation we did not have time allocated to address.

One of the most important features in our future work is to add players for comparison. As previously mentioned, this will require further architecture and development to create reliable and efficient services that deliver data to the user asynchronously. This would also require further user testing to evaluate the effectiveness of visualizations when player comparisons are involved.

We also intend to add more toggles to our visualizations. For example, being able to toggle on and off the projected points vs. actual. This will give users more interactivity to visualize the exact data they need to make a decision. Additionally, adding toggles that support cross-visualization interactivity will be helpful so a user does not have to select each toggle individually. Making color improvements is also on our roadmap. We would like for the color palette to be minimal, color blind friendly, and to provide a consistent look and feel.

Each of us has performed a systematic review of papers exploring topics of sports visualizations. With the research we have conducted, we would like to integrate key ideas that support tasks of analysts and sports fans. Through this process, we discovered that visualizations in sports has had a limited impact on fantasy players, amateur data consumers, and fans. There may be some applicable lessons from research, but we are mostly working in uncharted territory. This may include the incorporation of dashboards, applications of computer vision, or new data collection methods.

VI. Conclusions

The goal of our project is to make fantasy football more accessible to new users while providing novelty and increased judgment to veteran players when having to consider player transactions. We developed four visualizations: a projection evaluation visualization using a line char idiom, a past performance breakdown using a stacked bar chart idiom, a team fantasy scoring visualization using a treemap idiom, and a depth and injury chart using a custom idiom. The combination of our visualizations were used to aid fantasy football users in making an informed decision about which players will perform the best on a given week and should therefore be added to their team. The effectiveness and usability of our visualizations were evaluated using a case study, where 14 fantasy football users were provided a specific scenario and task. Overall, user feedback was positive and suggests that our visualizations were sufficient in answering the provided questions. Although, some users thought that our depth and injury visualization was difficult to understand at first glance.

Using the feedback received, we first want to improve our current visualizations. This includes adding notes on how

to use our visualizations effectively and creating a cohesive dashboard for our users. We would also like to extend our work to allow for comparisons between at least two players. This will require further development and user testing to evaluate the effectiveness in making comparisons in addition to conveying player information. However, if we can extend our framework to compare players then it would be a much more valuable tool for our users in the future.

REFERENCES

Note: We were not able to find applicable references. In the authors experience during the systematic review process, there is a lack of research and writing specific to the field of fantasy football and data visualizations.