

Basic Info

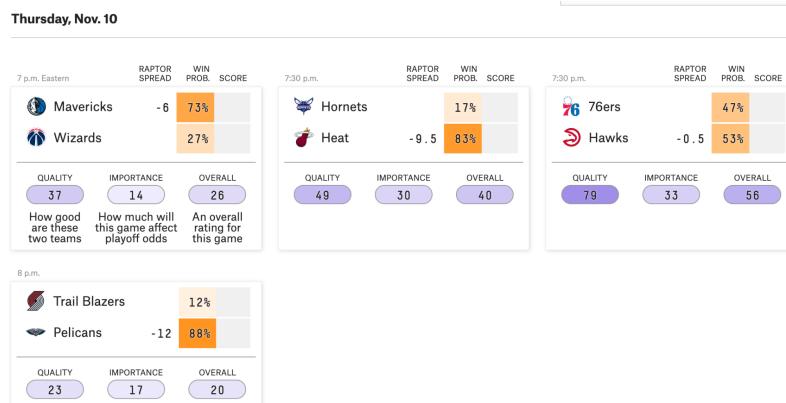
- **Project Title:** BeatingTheOdds
- **Team Members:**
 - Archie Menon, u0928536@utah.edu, u0928536
 - Mark Patterson, u1207184@utah.edu, u1207184
 - Tyler Gaul, u1195904@utah.edu, u1195904
- **Repository Link:** <https://github.com/Wilferd/BeatingTheOdds>

Overview and Motivation

There's a saying when it comes to gambling that "the house always wins." The motivation for our project is knowing how often Vegas predictions are right for the outcomes of NBA games. We are interested in sports betting and want to display the results for the NBA and compare actual outcomes to predicted outcomes. For every NBA game, there are countless statistics that oddsmakers come up with such as money lines, spreads, favorites, expected payouts, etc. Our goal is to distill this information into helpful visualizations that show how teams perform vs how they are expected to perform.

Related Work

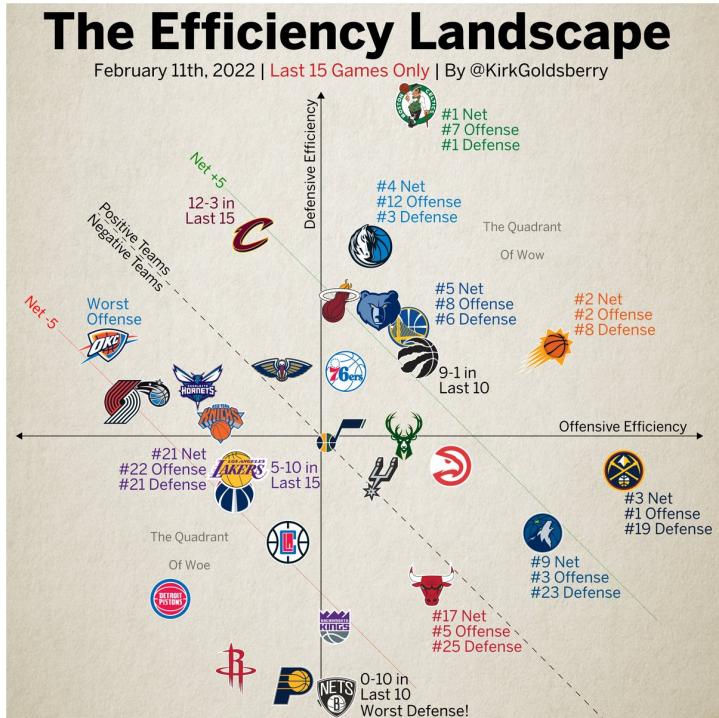
While not specifically related to Vegas gambling, NBA odds predictions have been done (and visualized) by the website FiveThirtyEight which is very prominent in sports analytics.



For all NBA games, FiveThirtyEight has a prediction on which team will win and by how much. We were interested in how good these sorts of predictions are. FiveThirtyEight is a lot more open about their methodologies, but there is no doubt that Vegas also employs mathematical models to predict the

outcomes of NBA games. Visualizing predictions versus actual outcomes is what we found to be an interesting topic for our project.

Kirk Goldsberry of ESPN who has also worked at FiveThirtyEight has also been a source of inspiration for visualizing different aspects of the NBA. Here's an example of visualizing the different offensive and defensive ratings of the NBA:



Our project employs a similar strategy of using NBA team logos within visualization.

Questions

The questions that we are trying to answer are:

- How often is Vegas right in predicting bets?
- Is winning and beating the spread related?
- What teams usually beat the spread?
- Do certain matchups have a consistent result in regards to the spread being beaten?
- Which of the five betting websites predict the correct result the most often?

The main question was how good is Vegas at predicting bets. These questions evolved to include how specific betting websites performed and which teams are predictable and which aren't. The new questions we considered were about how teams perform in certain matchups and overall how good a team was in a season.

Data

Source: <https://www.kaggle.com/datasets/erichqiu/nba-odds-and-scores>

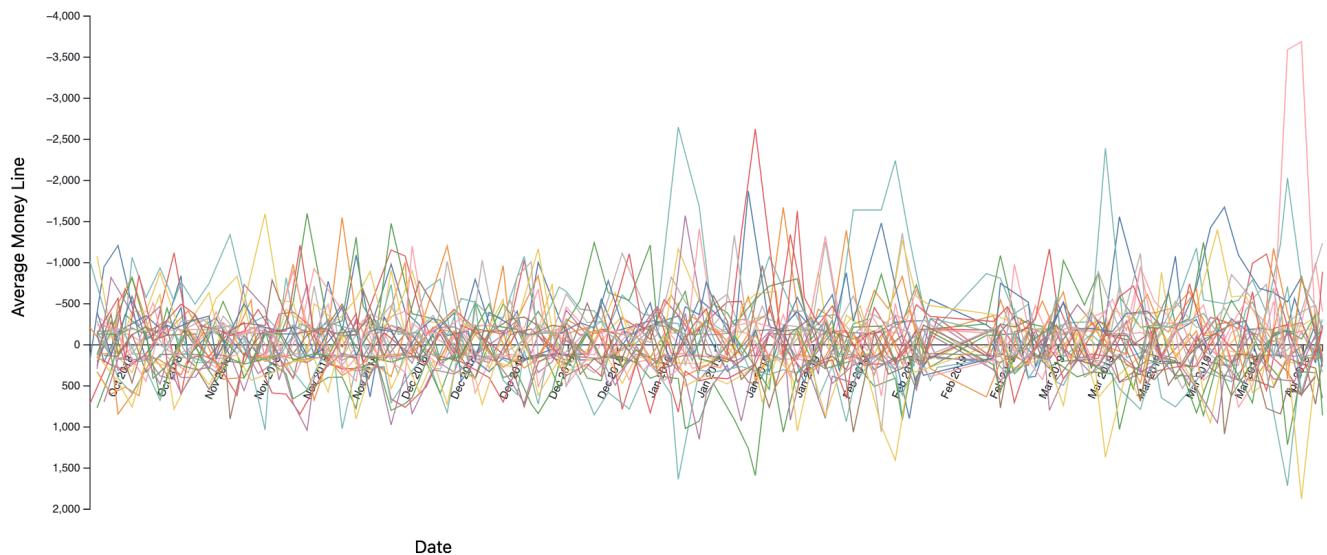
- The source we used is kaggle.com. We found a data set containing NBA odds and scores from the 2012-2013 season to the 2018-2019 season. This data includes the results of each game in the season, the predicted spreads from multiple betting websites, and the best spread choice.

Exploratory Data Analysis

The first iteration was very similar to our project proposal but the main issue we were facing was too much data on the line chart at once. Since there are 30 teams and 82 games, we found that displaying that many data points at once is really confusing for someone trying to deduce anything from the visualization. We tried basic ways of trying to fix this such as making the svg larger or removing lines but ultimately we decided that the best way of displaying our data will not include all of the teams plotted at once.

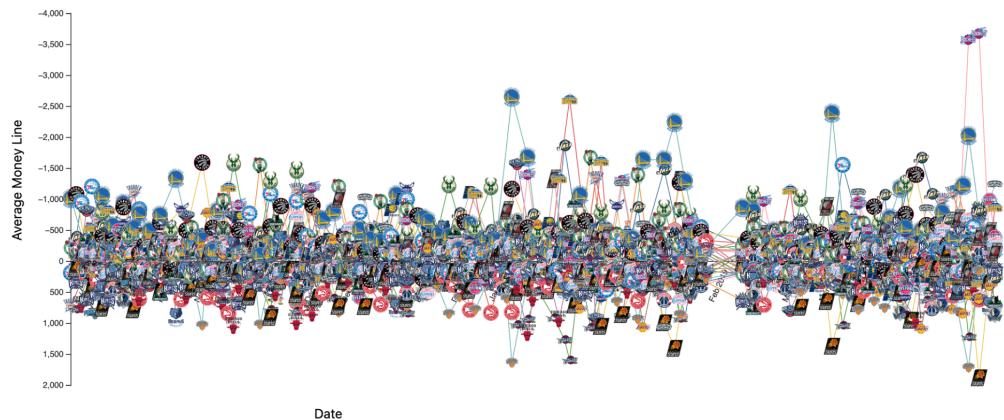
Initial Line Chart

Beat the odds



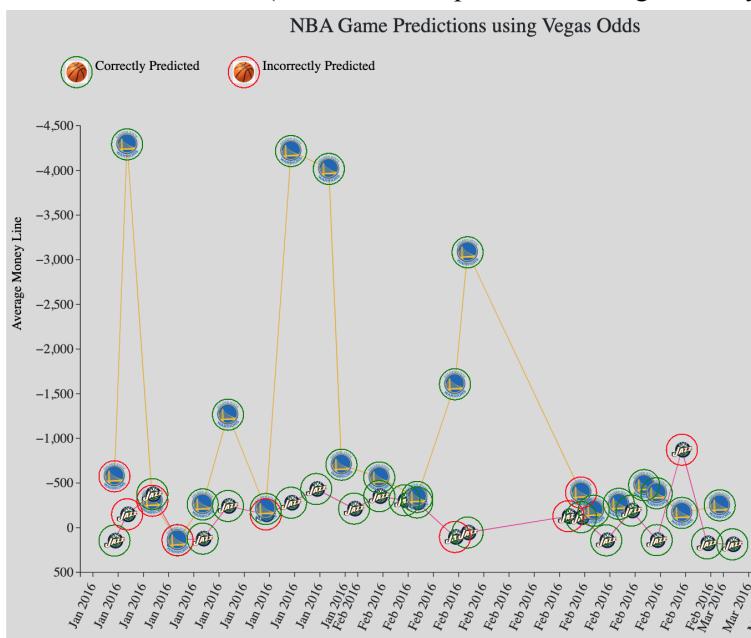
Initial Scatterplot

Beat the odds



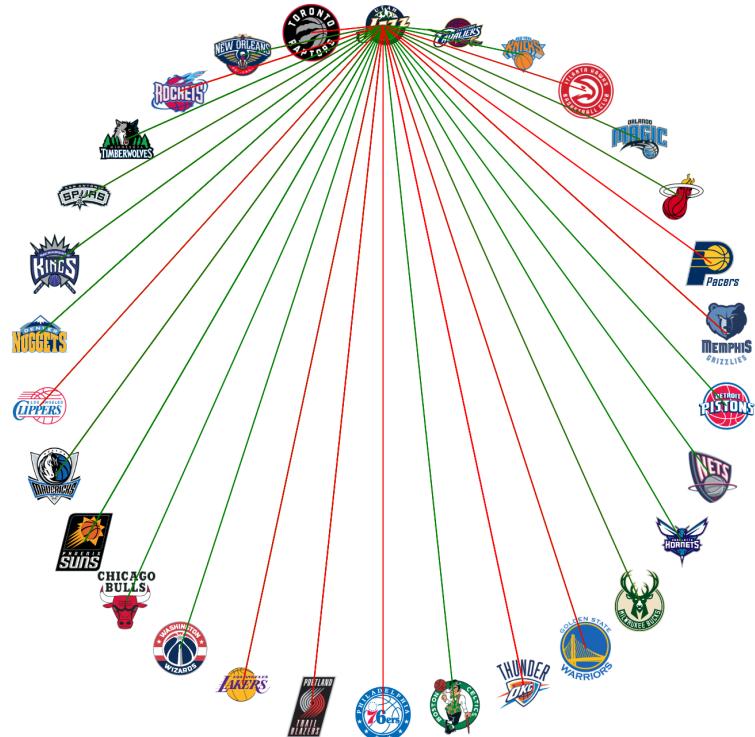
These initial plots show how big the dataset is for a singular NBA season. Each of the 30 teams has 82 games which is a lot to visualize. Our design for the line chart remedies this by allowing a plethora of filter options including: a toggle button for every team, date filtering, a toggle button to show the most mispredicted team as well as the most correctly predicted team.

Here's a more filtered (2 teams and a specific date range of early 2016) version of the line chart:



The first version of the chord diagram was similar to how we structured it in the proposal and it helped us visualize different teams head-to-head matchup data very easily. The issue we found though was that it was difficult to decipher the teams that were close to the top of the circle leading us to seek a different arrangement of the data.

Initial Spoke Diagram



Spoke Diagram with Centered Selection



This is the design that we decided on as it is instantly easy to see a trend with the team that is selected and all of the teams are equally visible.

Design Evolution

There were a couple different visualizations that we initially considered. One that we knew that we were going to have was a line chart because we knew that this would be the cleanest way to display all of our data for money lines. The issue that we instantly faced with this was that there was so much data for a season. To remedy this solution we initially in our proposal decided that we were going to use a brush so that we can filter different parts of the season and make it less of a hairball. Eventually, we decided to just use a date picker as that way the user has more fine grained picking of the data range.

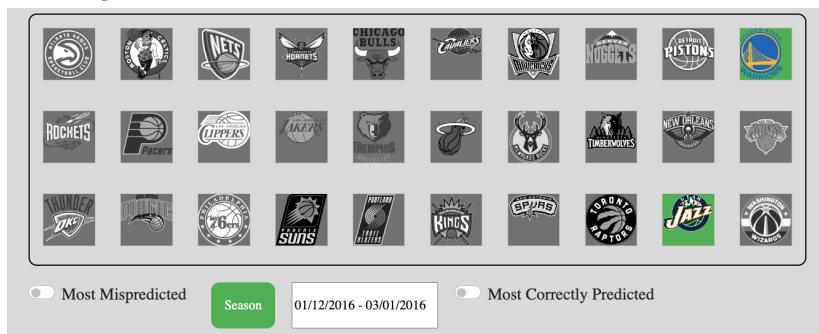
Another visualization that we considered was a bar chart to show the differences for the different betting sites. We decided that it would be best if the main line chart showed the average money line and then hovering the bar chart to the side of the line chart would display information about that specific game and for all of the different betting sites. The bar chart overall stayed basically the same as when we initially designed it in our five design sheets that we had in the project proposal.

Lastly, we had the chord diagram that we wanted to show head-to-head matchups. We thought this would be insightful as it is hard to really see how a single team is doing specifically in the main line chart. This would help see which teams covered the spread the most and won the most in a single interaction. We then saw that having them all in one circle made the visualization harder to read so we decided to change it to a spoke diagram. This way the team that was selected was not only in the center and easy to catch with the eye but it also allowed for nicer transitions when selecting different teams.

Overall, we mostly stuck to our initial proposal and improvised when needed. Most of the changes were not super major as our initial proposal nearly mirrors our final design with some nice quality of life additions.

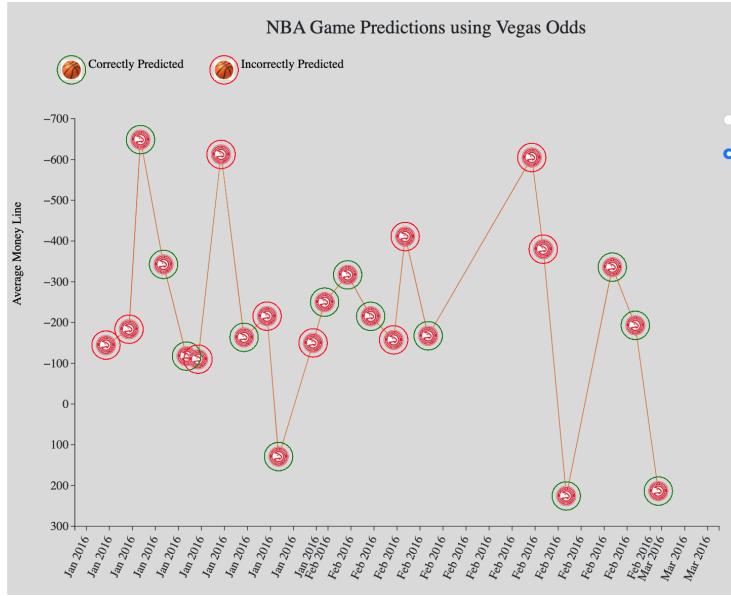
Implementation

Filtering

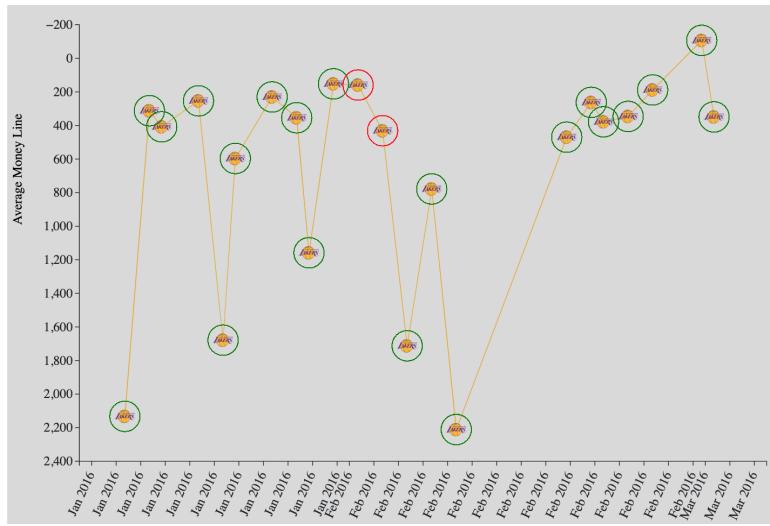


At the very top of the page, the line chart has a lot of filtering options. You can select multiple teams to compare (currently the Golden State Warriors and the Utah Jazz are selected). Selecting these updates the line chart to show those teams. The Most Mispredicted and the Most Correctly Predicted switches will show the team that fits the relevant criteria during the selected date range.

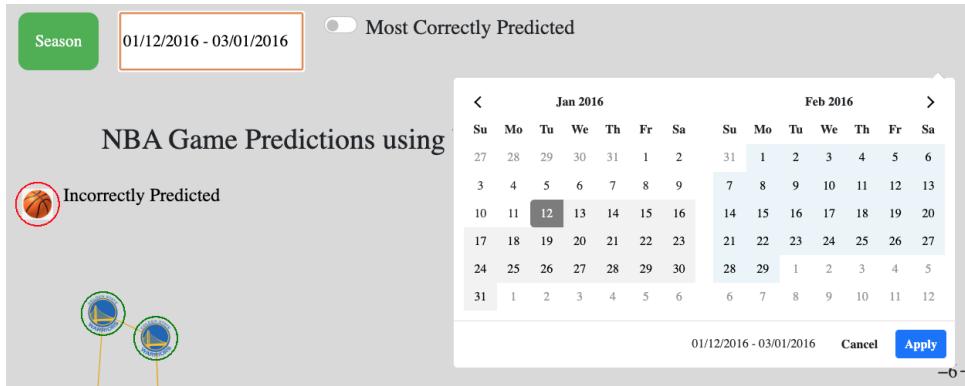
The most mispredicted team in this range is the Atlanta Hawks:



The most correctly predicted team in this range is the Los Angeles Lakers:

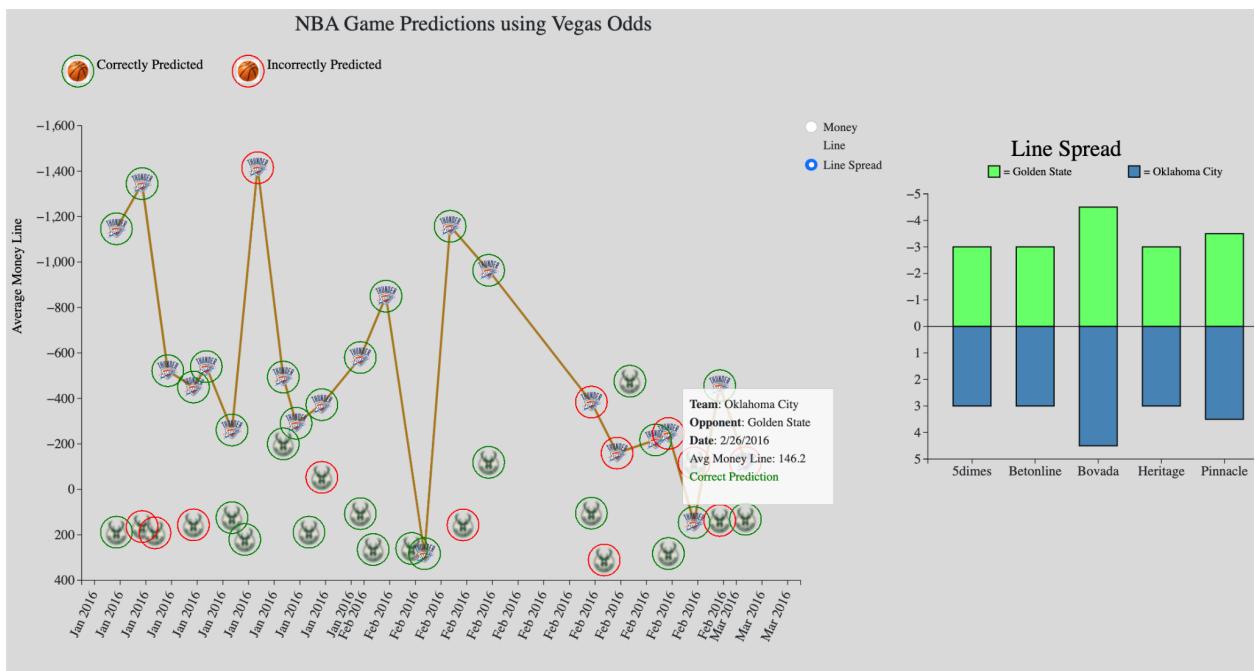


There is also date filtering which is done through clicking the text box and shows a calendar popup:



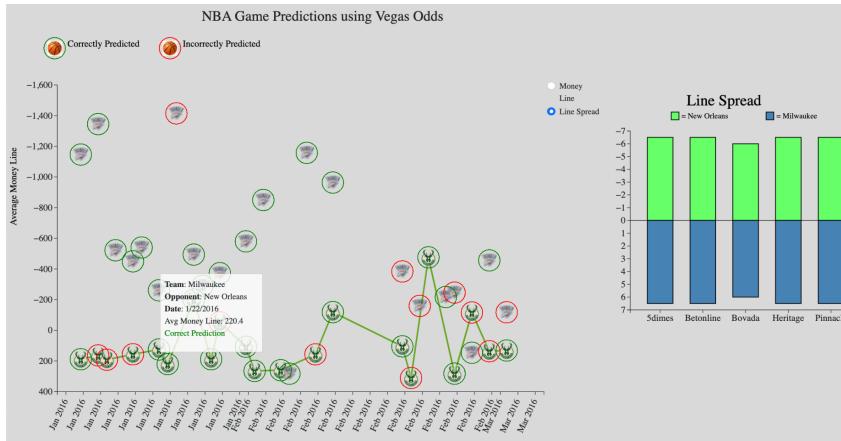
This date filtering is bounded by which season is currently selected. If you want to see a different NBA season, you click the Season button.

Line Chart and Bar Chart



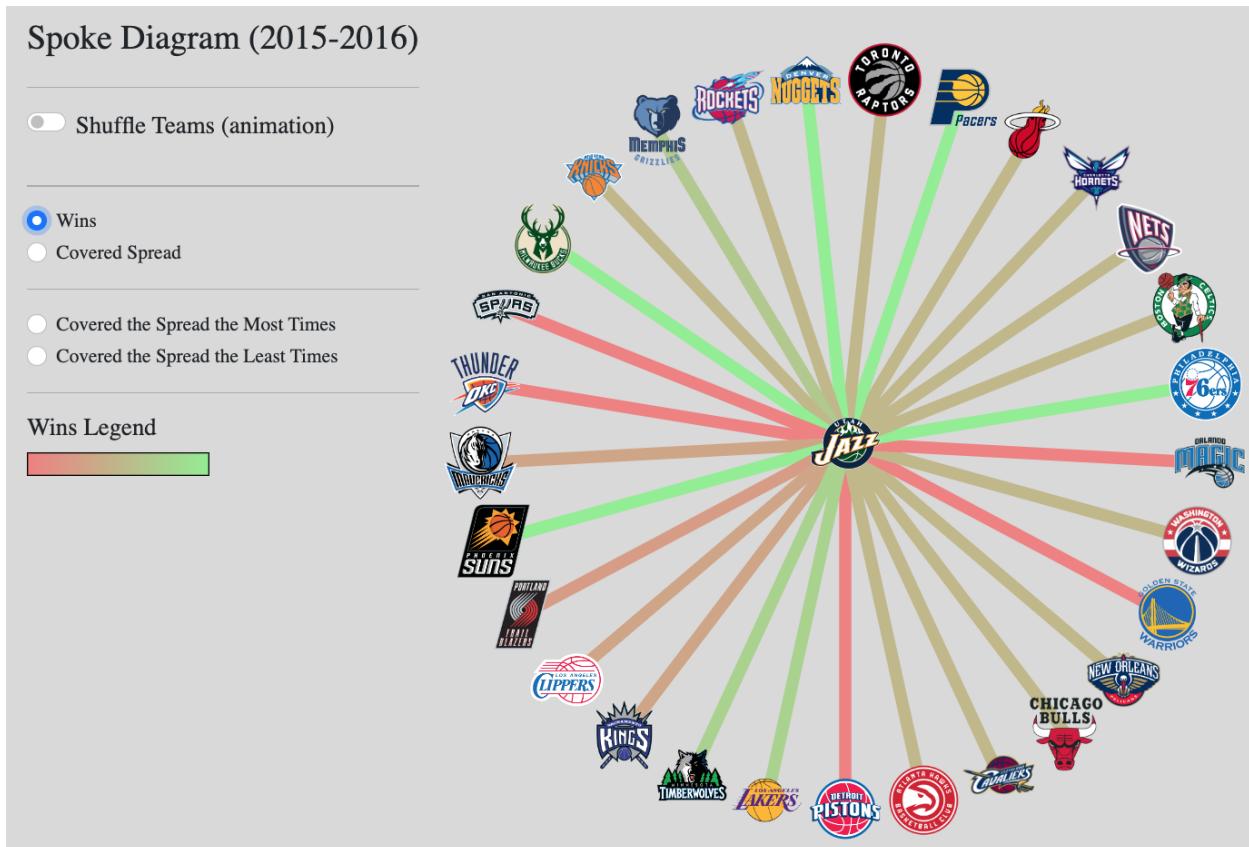
There is a line chart and a bar chart that are interactively connected. The line chart shows the aforementioned filtered data. Each point of the line chart is an NBA regular season game. The X-axis is time and the Y-axis is the money line (how favored or unfavored a team is to win and negative means favored). Using Delauney triangulation and Voronoi diagrams, the closest line to the mouse is put into focus and a tooltip appears above the closest point showing details about the game.

On the right hand side, there is a bar chart that updates depending on what game the mouse is closest to. This bar chart shows the predictions for the individual sportsbook in the dataset while the line chart shows the average. The bar chart also has radio buttons on the left side to switch between the aforementioned money line or the spread. The spread is how much a team is expected to win or lose by.



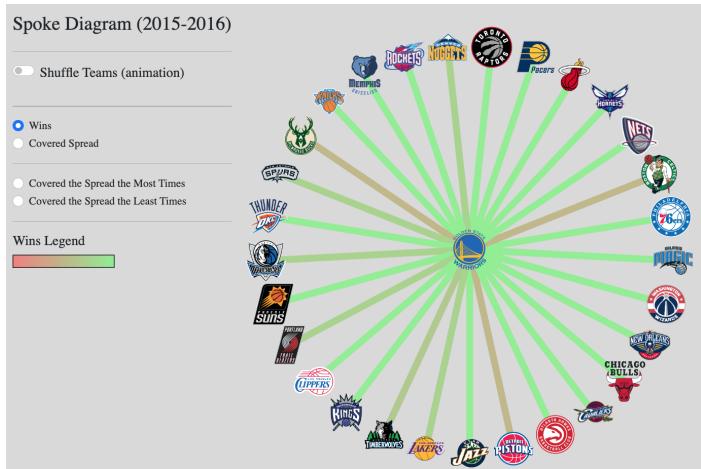
The above image just shows that a different mouse position shows a different line, game, and bar chart.

Spoke Diagram



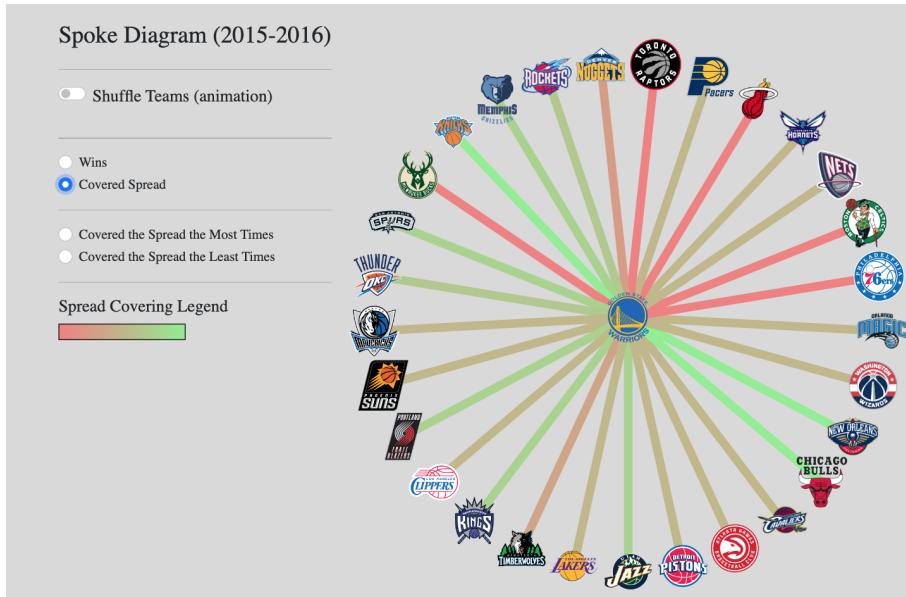
This spoke diagram is for visualizing the head to head data between different teams for the season selected from the aforementioned Season button. In the above image, the Utah Jazz are in the middle and have a lot of wins that are shown in green and some losses that are shown in red. Since you play multiple teams within a season the color is interpolated between red and green, so if 2 teams have won an equal amount against each other, the line between them is a beige color.

Clicking on another team on the outside of the circle changes the spoke diagram. For example clicking on the Golden State Warriors shows:



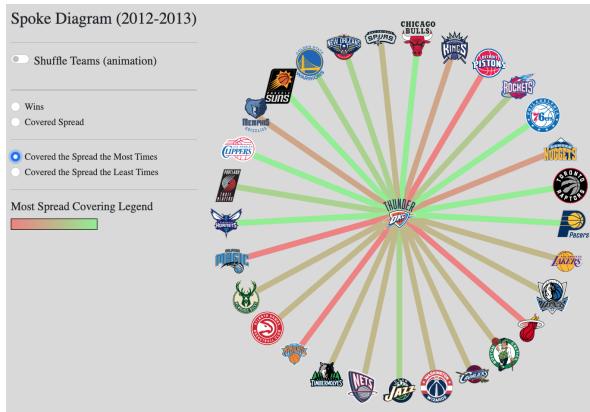
This season the Warriors were especially good so they have a lot of green lines. When selecting a new team a cool animation is shown, and this animation could be changed to a shuffle.

Clicking the Covered Spread option updates the diagram:

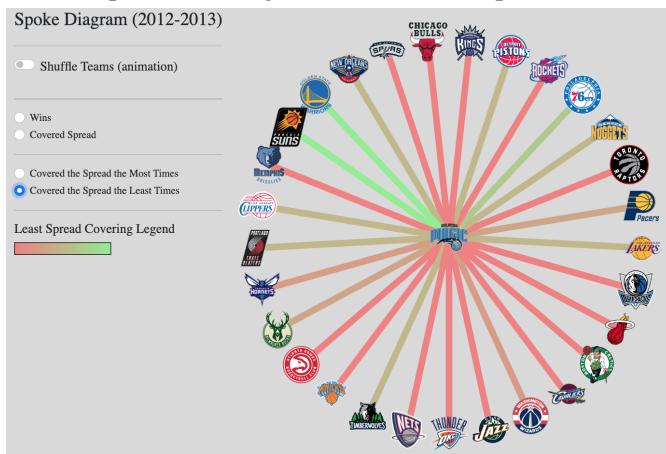


Covering the spread means if a team outperforms the expectations for how much they'd win or lose by. For example if a team was predicted to lose by 3.5 points, they could either win or lose by less than 3.5 to cover the spread. In this example, even though the Warriors won a lot of games as shown earlier they didn't cover the spread nearly as much (which is to be expected).

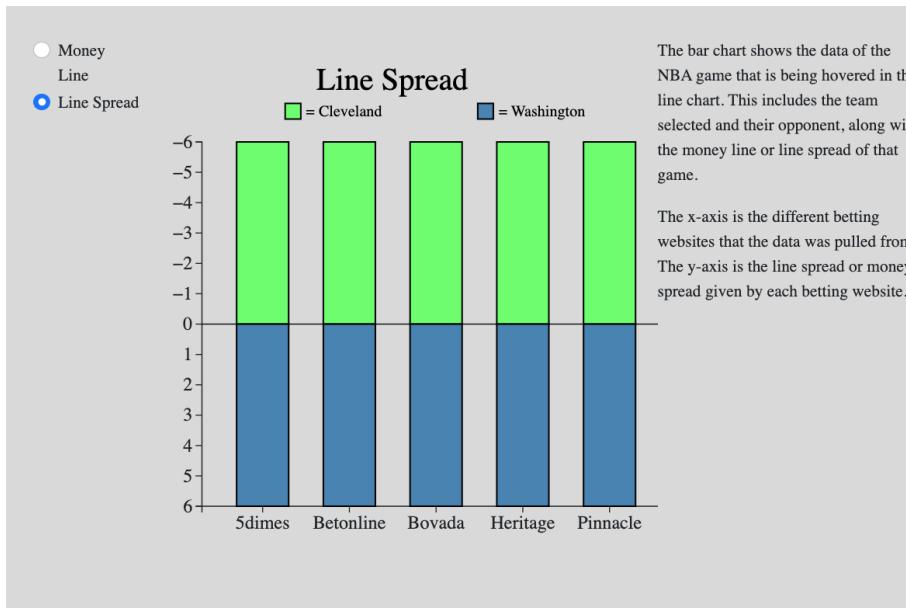
There is also an option to show who covered the spread the most amount of times:

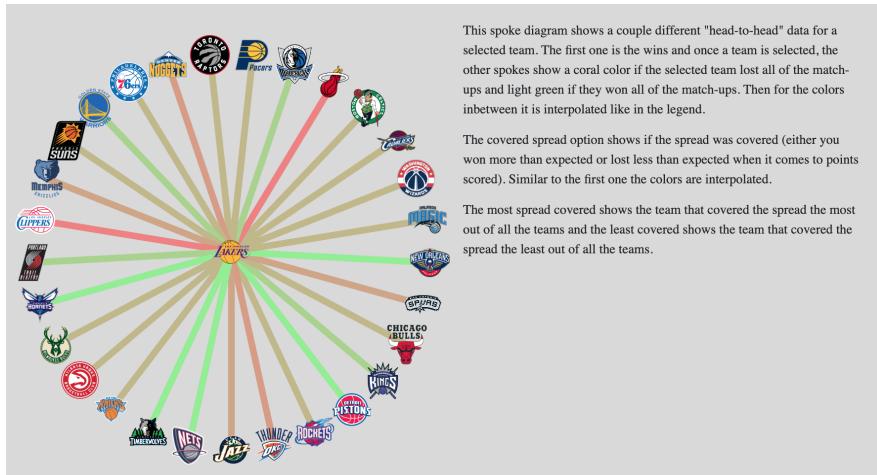


And an option showing who covered the spread the least amount of times:



Our implementation also has helpful descriptions next to the bar chart and the spoke diagram:





Evaluation

We learned a lot by looking at our visualization. We learned how often Vegas correctly predicts a game's outcome. We also learned what teams performed as expected and what teams didn't, along with what teams were good and what teams usually covered the spread. Another thing we learned is that the predictions of each website were similar, but sometimes one website would be a little different.

We answered the question of how often is Vegas right in predicting bets using the line chart with green or red circles showing if the prediction was right. We answered the questions, is winning and beating the spread related, and what teams usually beat the spread, with the spoke diagram that shows how much the teams win and how often they covered the spread. We answered the question of whether certain matchups have a consistent result in regards to the spread being beaten with the spoke diagram as well. We answered the question of which of the five betting websites predicted the correct result the most often with the bar chart that shows what each of the five websites predicted.

The visualization works very well in displaying the data and allowing different conclusions to be drawn. The line chart, bar chart, and spoke diagram all work together to show how well Vegas makes predictions and what the actual results were. One possible improvement could be showing the actual score of the selected game to show how close or far the line spread prediction was. Another improvement could be specifically showing which betting website had the best predictions and which had the worst.