Visualizing Consensus

By Hilson and Kartik

In this visualization, we attempt to explain the level of agreement or disagreement in a consensus ranking, and eventually compare between multiple consensus rankings.

Overview:

Consensus ranking is essentially a way to find a ranking which “best” represents the individual contributing rankings. Generation of consensus rankings may involve algorithms. Simpler algorithms can even be as simple as aggregation of individual rankings. It is important to understand what “best” means. A quick measure of “best” could be exploring how much the consensus ranking agrees or disagrees with individual rankings.

Parallel Coordinates Plot (PCP) is a great choice when it comes to visualizing multivariate numerical data and it can be used to visualize such ranking scenarios which involve group-decision making where decision-makers build a consensus ranking that “best” represents the collection of base rankings.

Related Work:

Literature:

Hindalong et. al[3] used the PCP technique to visualize the ranking problem since it helps users to visualize how individual candidates are ranked by different rankers. However, to visualize how a consensus is reached is not possible using just PCP because of a limitation: Comparison is limited to adjacent axes. To solve this problem, the system integrates box plots and strip plots for the analysis of consensus ranking. However, the system was designed for a limited number of rankers and candidates and does not scale up for a large dataset.

Another approach to visualizing consensus is to use color-coded stacked histograms (J Bok et. al [1]). While this technique allows the comparison of non-adjacent axes and could be effectively used to compare consensus ranking with all the base rankings, the system would not scale up for a large dataset with a large number of rankers. Users will have to scroll through all the base rankings to see how consensus aligns with it.

To solve this, we attempted to integrate parallel coordinate plots with idioms like box plots or strip plots. We hoped that our design would make it possible to analyze the degree of agreement of base rankings with consensus ranking. Furthermore, to visualize and inspect a large dataset, we attempt to exploit the spatial distortion technique (Sarkar et. al[4]).

Initial Design Idea:

After going through the literature, we identified a common limitation in all the existing methods: scalability to large datasets. We set out to explore ways of extending any method to large datasets.

At this point one of the first designs was sketched out. One of the first drafts of this visualization looked like:

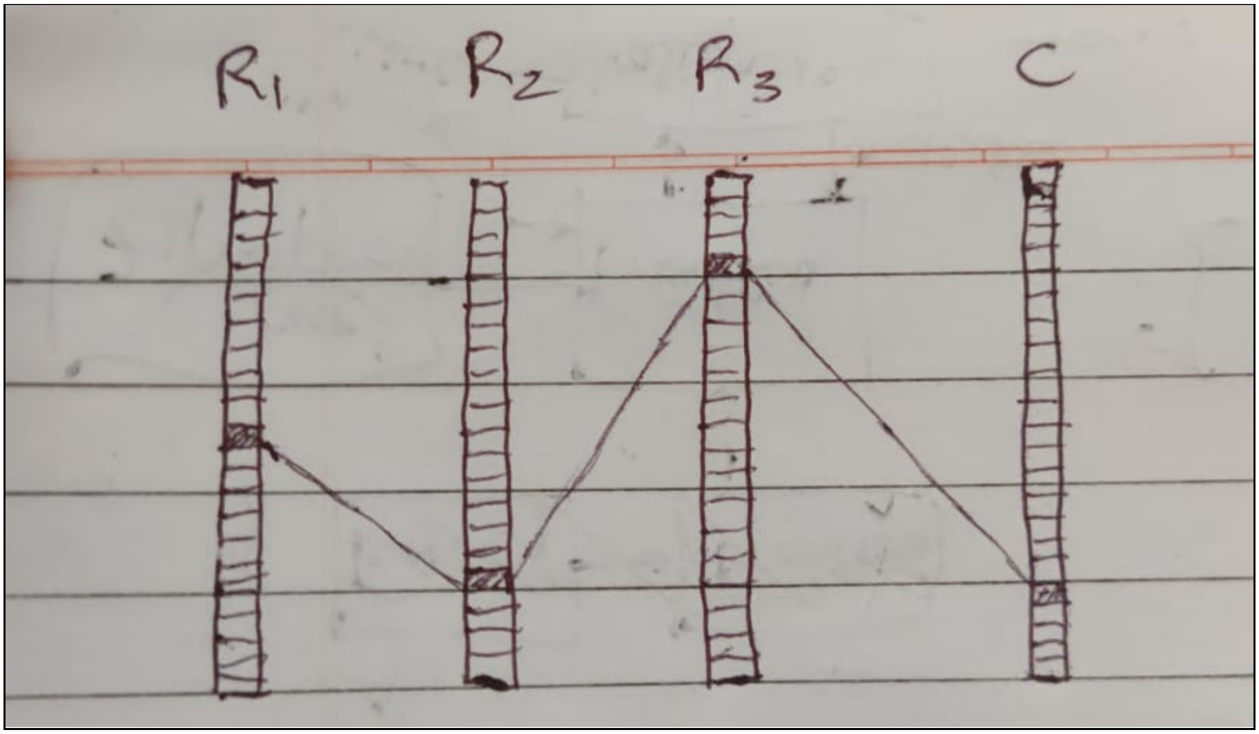

Fig 1. Initial design sketch

But it was quickly realized that it was not possible to scale this up. If the number of candidates increased in the visualization, we would need to scroll through to find a particular candidate in the rankings which was not found to be ideal. If we were to fit all the candidates on one page, the size of each candidate would decrease so much that we would not be able to see individual candidates.

It was essential that the user should not have to scroll through to find the ranker. Hence, the size of individual candidates was decreased to fit the page. A modification to the previous design was considered which would involve creating a hover feature that would allow the user to hover over the part the user is interested in and let the user see a zoomed version of it using the fisheye distortion effect.

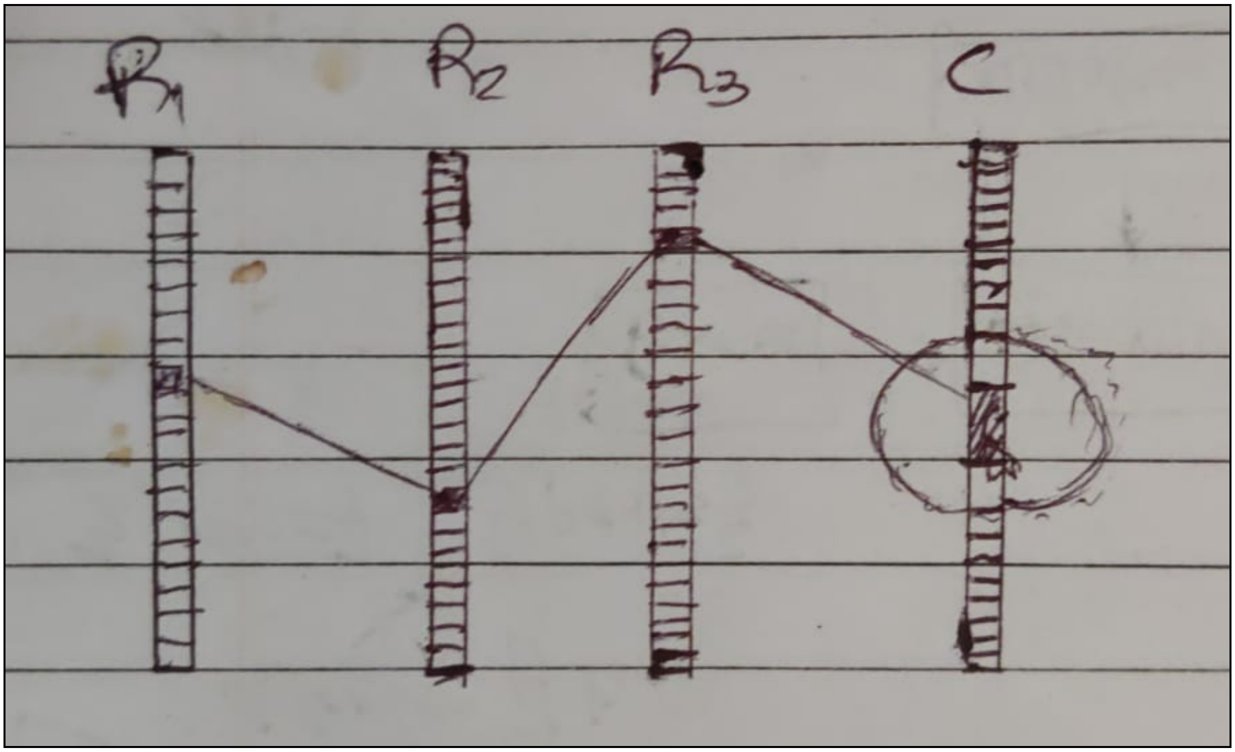

Fig 2: Attempt fisheye view as in https://bost.ocks.org/mike/fisheye/

We explored a few implementations of the fisheye distortion effect and this website captured the effect perfectly. The first plot on the website implements the effect on a network. It further shows us how the effect would look on a grid in the second figure. A few limitations of the fisheye as identified are that it compresses data instead of zooming into it near the circumference of the fisheye and that it causes distortions.

Although these limitations do not affect our design, another approach in this area is a cartesian distortion. This ensures that there is no distortion. This approach was explored as well, but just from an aesthetic standpoint we decided to go with fisheye distortion.

Motivation:

So after reviewing some literature and coming up with the raw sketches, we decided to implement something on the lines of Fig 2. It was felt that this would explain the consensus well and could be scaled to large datasets as well.

Our approach started with addressing the problems stated for large datasets. As we went about understanding the existing methods, we understood that our method is able to answer an important problem which existing methods cannot: understanding level of agreement or disagreement in the consensus when compared to the original contributing ranks.

We thought that we addressed the problem of large datasets by considering the fish eye distortion. This led to a change in the direction of our project. We started looking at ways in which we could address the problem of showing the agreement or disagreement in a better way instead of just showing lines connecting to the points in individual rankings. We were able to brainstorm and think of multiple methods to show the agreement or disagreement of individual elements in a consensus by focusing on the difference in the rank between the consensus and the individual ranks.

While explaining this method of visualization, we found an obvious problem we did not think of: How can a person seeing our method for the first time understand what is happening in the visualization?

We shifted focus from trying to create these visualizations for different datasets to creating a method which would allow a person to understand the visualization better and help interpret it. A static approach was first considered to show how we end up with our visualization. But, that would not comprehensively explain the method.

An interactive approach was then considered for the same. Scrollytelling was one way to showcase a step by step procedure. The questions we chose to answer now were based more towards the understanding aspect of things: How can a user best understand what is happening in the visualization without having to read any text?

Datasets:

For the demonstration of the project, we used following datasets:

1. 60 candidates data generated from http://roycekimmons.com/tools/generated_data/exams
2. 500 candidates data generated from https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

For consensus rankings:

1. Consensus rankings generated using FairCopeland Algorithm (Cachel et. al - ManiRank[6])

Exploratory Data Analysis:

What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?

Initially we looked into implementing something similar to Group Decision Making visualization by Hindalong et. al [3] with box plot and dot plots. However, for large datasets, dots could be overlaid on top of one another causing obstruction. Box plot on the other hand could miss out on the internal distribution of the individual rankings. So, we decided with overlapping lines on top of another by adding some transparency.

For representing the consensus across all rankings, we also considered diagonally representing the position as in figure 3. However, it was not easy to understand and we had to switch to a different representation.

Design Evolution:

The previous few sketches show how we intended to address large datasets. The next few sketches show the visualizations that we considered were to address the understandability aspect of things.

Some of the visualizations we considered to increase understanding were:

1. Creating a consensus line and then comparing individual candidates to the consensus line.

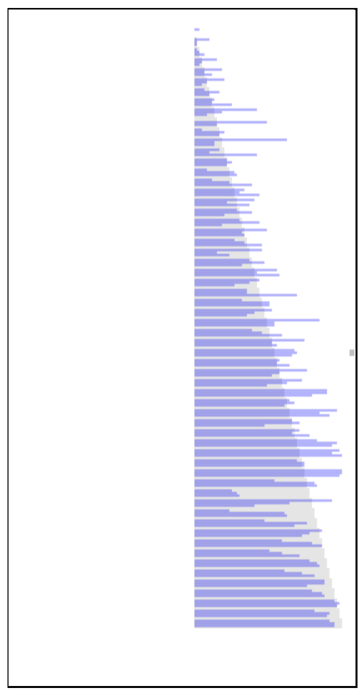

Fig 3: Diagonal consensus line

This design was considered but the issue was that the understandability was compromised here. This was also eventually not able to identify rankers.

2. Moving the axis to align the ranks with the consensus.

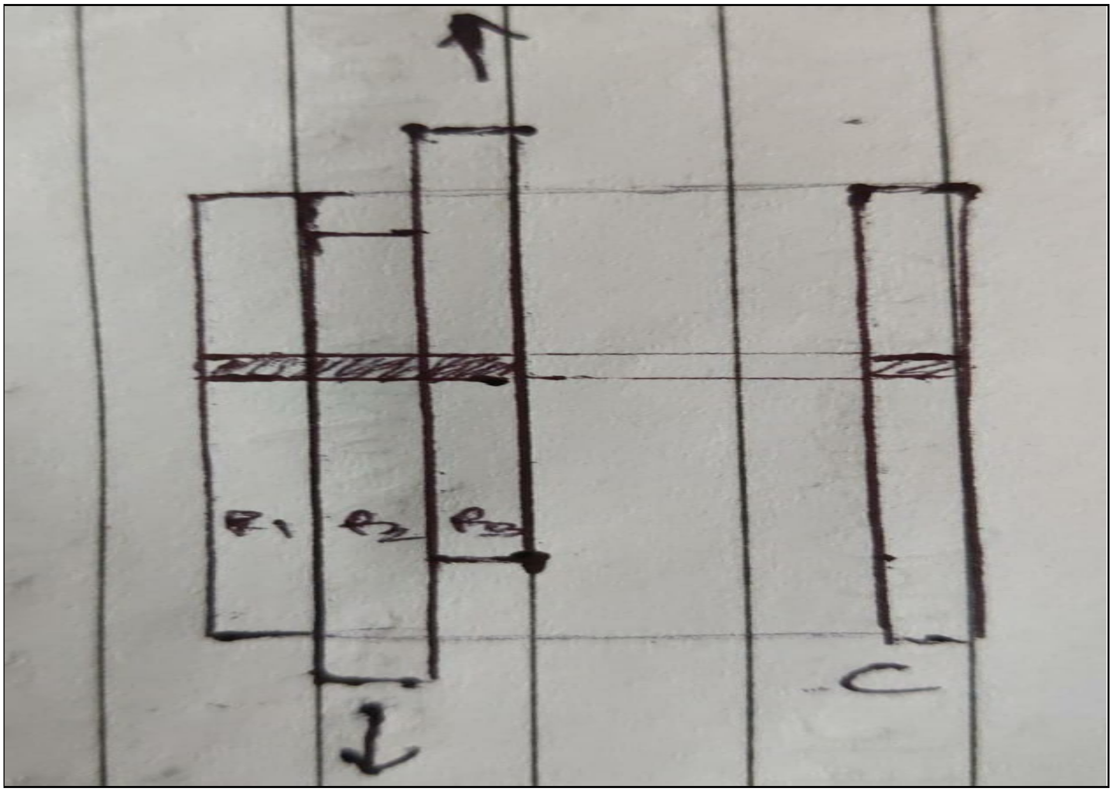

Fig 4: Offsetting axes to align a candidate

The offset of the individual axis would explain the level of agreement or disagreement.

3. Mapping the difference of individual rankings

   We realized that essentially we were looking at the offset to understand the agreement or disagreement which was the difference between the rank in the consensus and the individual ranks. We decided to map these differences because they directly explained what we needed and this is when we arrived at the following design.

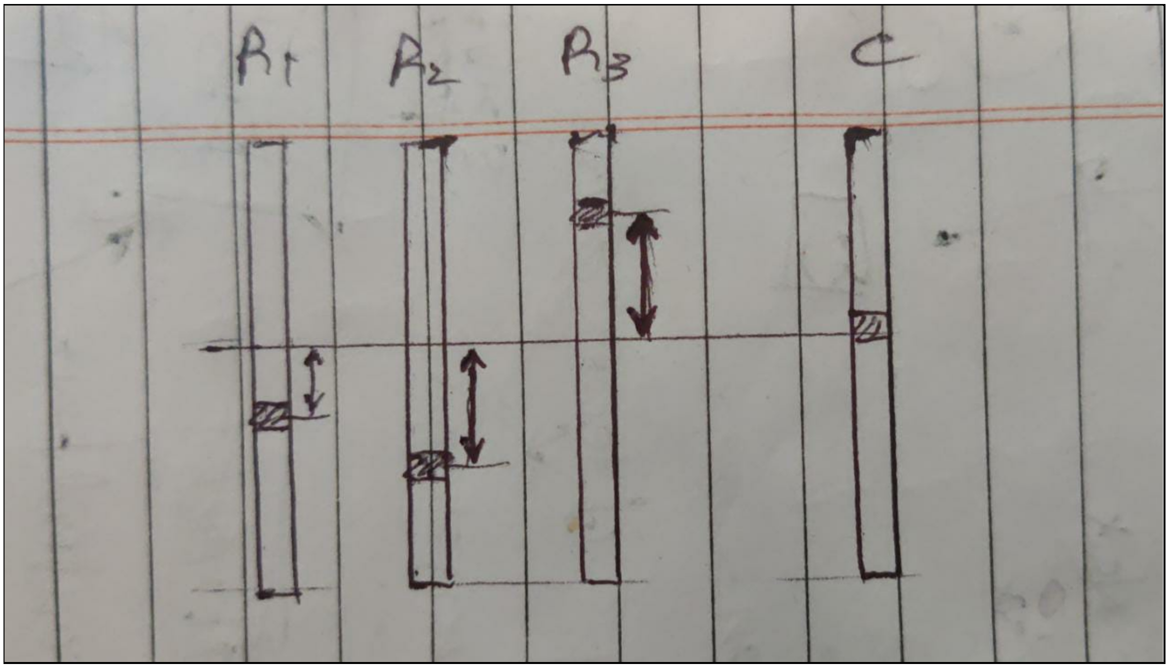

Fig 5: Difference in rankings

4. Mapping the differences

   The larger the difference the larger is the disagreement. But this was for one particular instance. Instead of hovering on the consensus we thought of plotting these differences for each of the candidates in the consensus ranking on a different axis as shown. This new axis was self explanatory in terms of showing the disagreement.

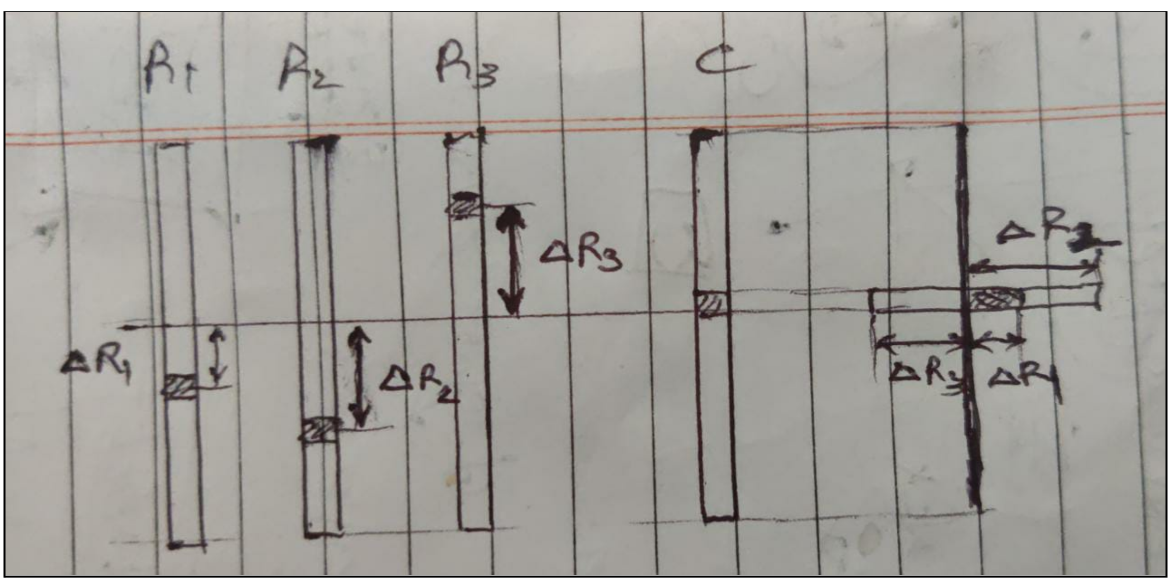

Fig 6: Mapping Difference in rankings

Implementation Goals:

The goals for our implementation were broadly:

1. To help understand how to read our visualization
2. To help understand the agreement or disagreement of the consensus rankings with the individual contributing rankings.

To achieve the first goal we have taken the help of a scrollytelling approach. The website loads where we are greeted with the message explaining what we are attempting to do. As we scroll down, the screen has 2 panels. The panel on the left explains the panel on the right which has the visualization. Following shows the first page of scrollytelling:

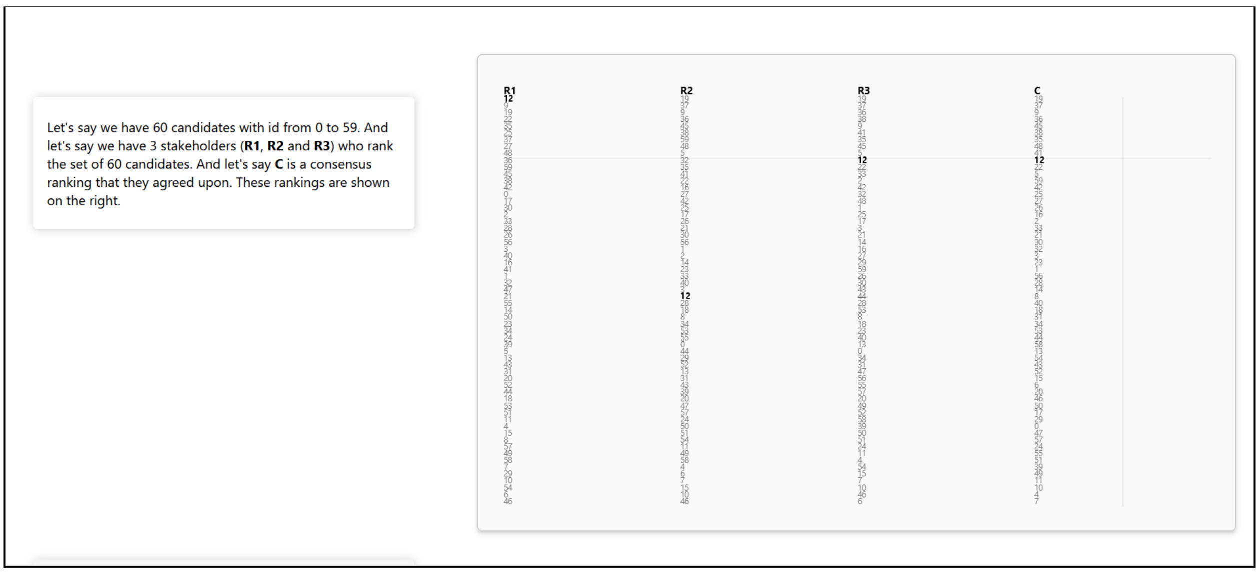

Fig 7: Scrollytelling page 1

As we go further down, the second page explains the differences or the disagreements among the individual rankers when compared to the consensus ranks as shown:

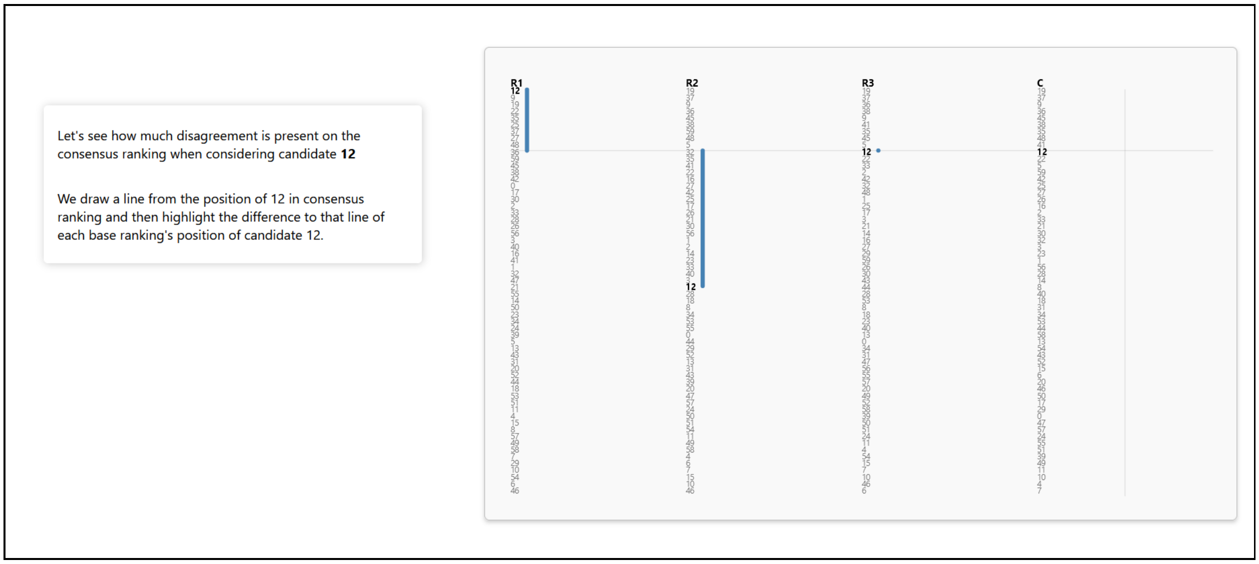

Fig 8: Scrollytelling page 2

These differences are then overlaid on the consensus axis to show the disagreements as shown:

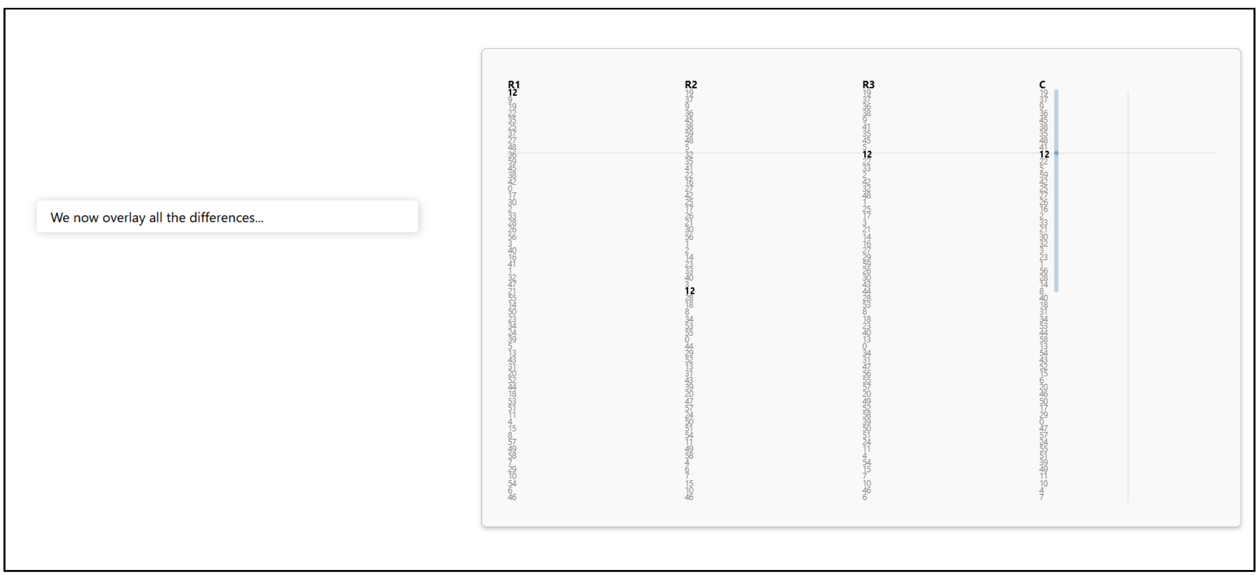

Fig 9: Scrollytelling page 3

Moving on, these disagreements are then rotated to show how we arrived at the final visualization.

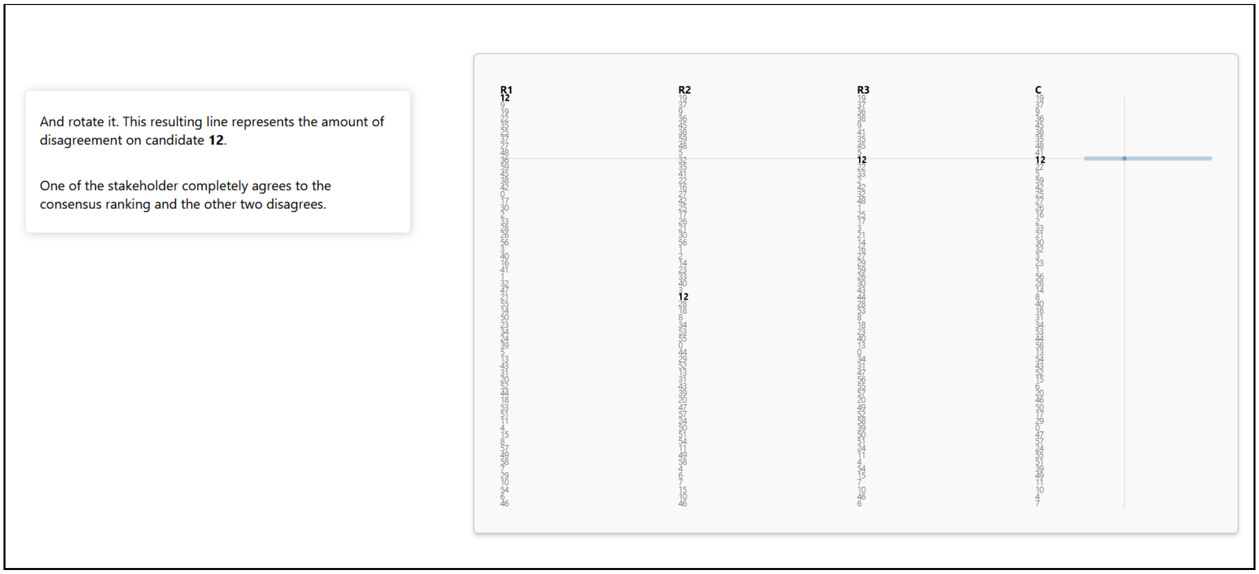

Fig 10: Scrollytelling page 4

Since this approach was shown for 1 particular candidate, the next panel goes on to show all the candidates in all individual axes

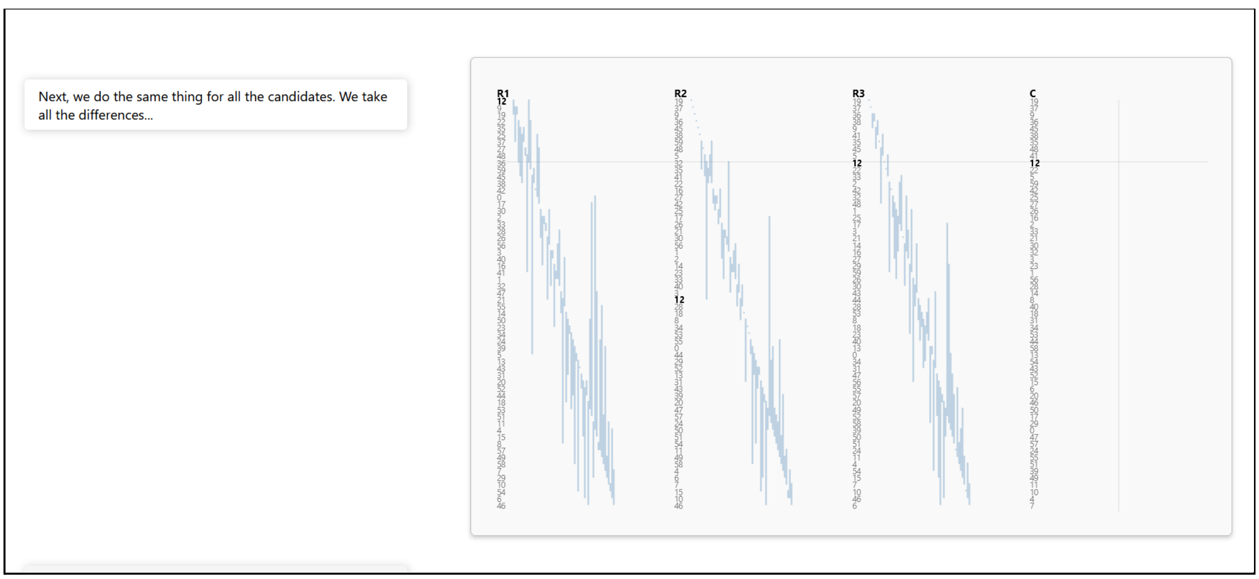

Fig 11: Scrollytelling page 5

These are then all overlaid together on the consensus axis as shown

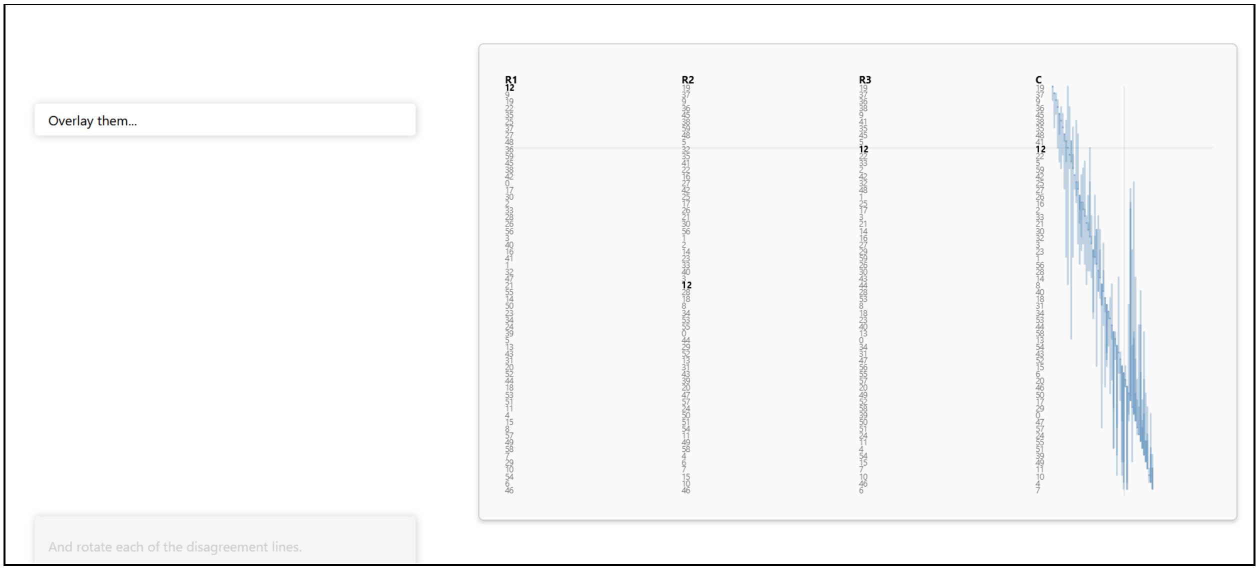

Fig 12: Scrollytelling page 6

They are similarly then animated to arrive at the final visualization for all the individual candidates.

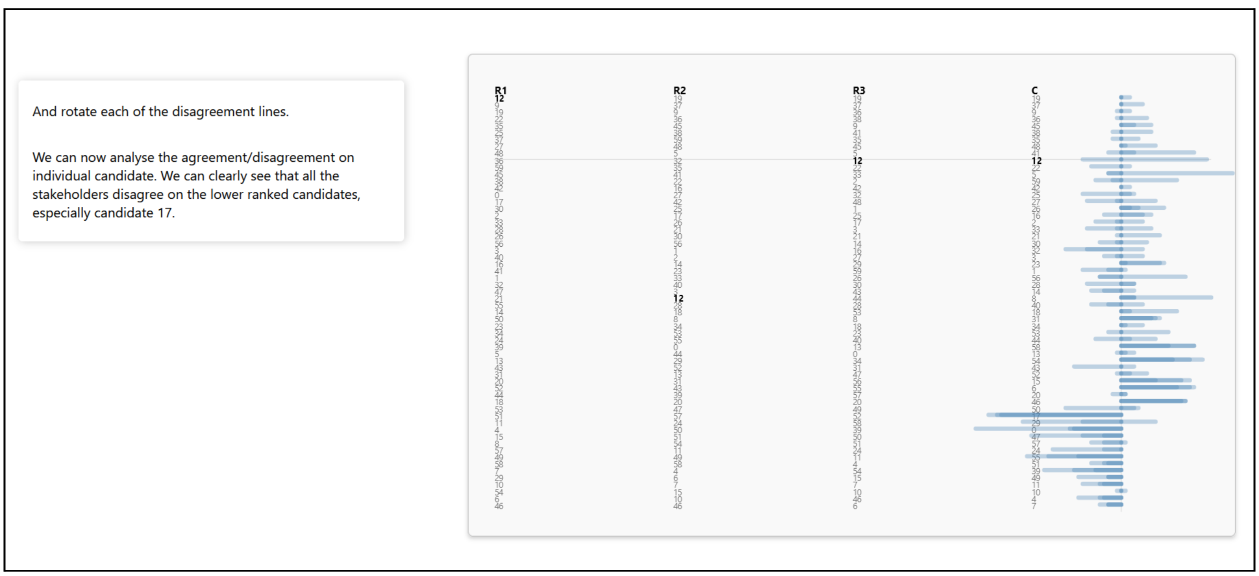

Fig 13: Scrollytelling page 7

Our next goal was to help the user understand the agreement or disagreement in the consensus with the individual rankings. After scrolling from the previous page, we land on to a page which has just one panel which shows 4 different consensus rankings as shown:

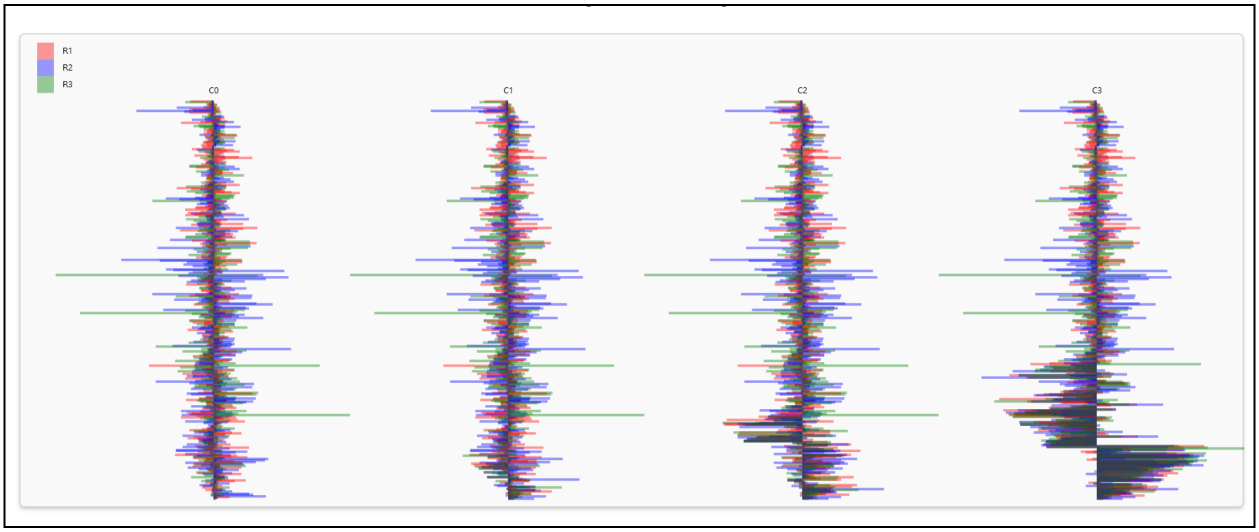

Fig 14: Final Panel

Clearly from the plot we can see that C3 has the maximum disagreement from the individual rankings, especially lower down the order as it has the maximum offsets from the axis.

Hovering over the legend allows us to filter the plot specifically for that particular individual ranking. This allows us to see the disagreement specifically for that axis. A sample is as shown for the R1 axis:

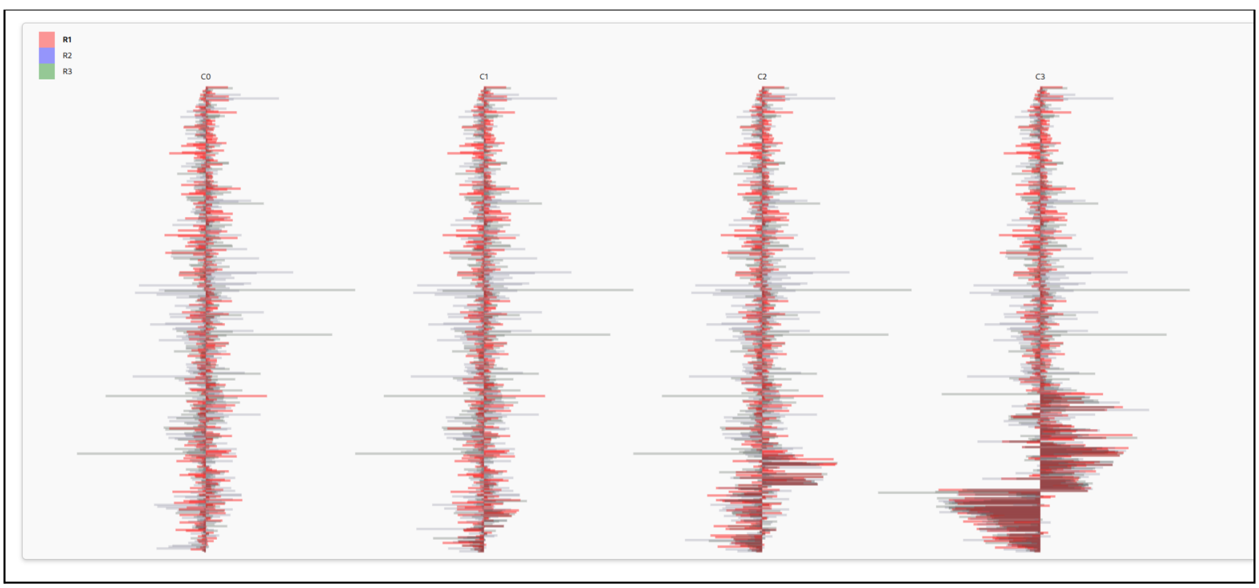

Fig 15: Final panel filtered by legend

Evaluation:

The questions we set out to answer are essentially answered using this visualization. Although it is not exactly how we started, we eventually modified our designs along the way. We are able to successfully answer our original question which was to understand the amount of agreement or disagreement each consensus has. We are even able to compare multiple consensus rankings.

We can probably make improvements on the usability of this visualization by creating hover features like fish eye view. Further a hover feature could also be extended to show the individual properties of that particular candidate.

References:

[1] J. Bok, B. Kim, and J. Seo. Augmenting parallel coordinates plots with color-coded stacked histograms. IEEE Transactions on Visualization and Computer Graphics, 2020.

[2] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. Lineup: Visual analysis of multi-attribute rankings. IEEE transactions on visualization and computer graphics, 19(12):2277–2286, 2013.

[3] E. Hindalong, J. Johnson, G. Carenini, and T. Munzner. Towards rigorously designed preference visualizations for group decision making. In 2020 IEEE Pacific Visualization Symposium (PacificVis), pages 181–190. IEEE, 2020.

[4] M. Sarkar, S. S. Snibbe, O. J. Tversky, and S. P. Reiss. Stretching the rubber sheet: A metaphor for viewing large layouts on small screens. In Proceedings of the 6th annual ACM symposium on User interface software and technology, pages 81–91, 1993.

[5] C. Tominski, J. Abello, F. Van Ham, and H. Schumann. Fisheye tree views and lenses for graph visualization. In Tenth International Conference on Information Visualisation (IV'06), pages 17–24. IEEE, 2006.

[6] K. Cachel, E. Rundensteiner, and L. Harrison. Mani-rank: Multiple attribute and intersectional group fairness for consensus ranking. In 2022 IEEE 38th International Conference on Data Engineering (ICDE). IEEE, 2022