

Prakhar Saxena
Professor Shetty
INFO 250 – 001
October 02, 2020

Project 1A

Introduction:

For this assignment, I chose the Orchestra dataset. This dataset has four attributes:

Orchestra type: String

Group: String

Instrument: String

Number: Number

The four charts that I chose to visualise this data are:

TreeMap

Circle Pack

Bar chart

Treemap

I experimented around with different attributes and patterns to produce the visualisations that provide the best understanding of the data as a whole.

Visualisations (In the next page)

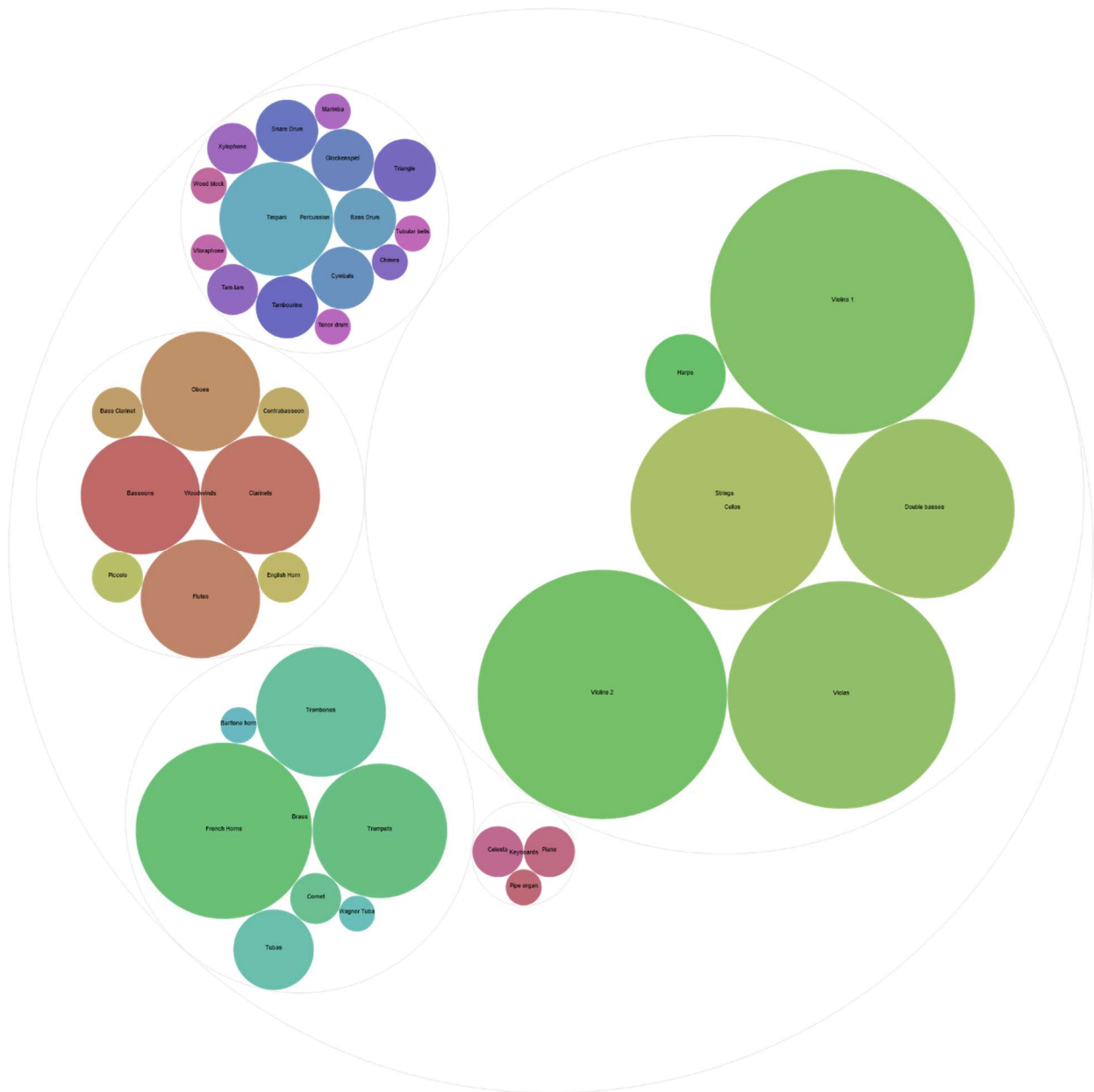
TreeMap



The TreeMap is a basic but a growingly popular visualisation used to depict the hierarchical shares of any data. I have used it in the past, to visualise the market share of the competitors of our business.

The treemap visualization supports four visual patterns: Hierarchy (numbers, strings, dates), Size (numbers), Color (numbers, strings, dates), and Label (numbers, strings, dates). Hierarchy is a required pattern and I mapped the dimensions Group and Instrument to it. For Size, I mapped Number, and for Color, I mapped Group. From the graph, we can draw easy conclusions that the green-yellow subset is the most predominant while the purple subset is the least represented in the Orchestra dataset. Based on the size, the instrument Violins 1 is the largest compared to any other instrument.

Circle Packing



This graph is similar to the Treemap, but a little simpler. It contains the same four visual patterns: Hierarchy, Size, Color, and Label. I mapped the dimensions of Group and Instrument to Hierarchy, and I mapped the Instrument dimension into both Color and Label, and finally, I mapped the Number dimension to size. Based on this circle packing visualization, the Strings group is the most represented in the Orchestras dataset and the Violins 1 subset is the largest in regards to size and number. The Keyboards group has the

smallest amount in both numbers of instruments and size. The Percussions section has the most diverse array of instruments.

Bar chart



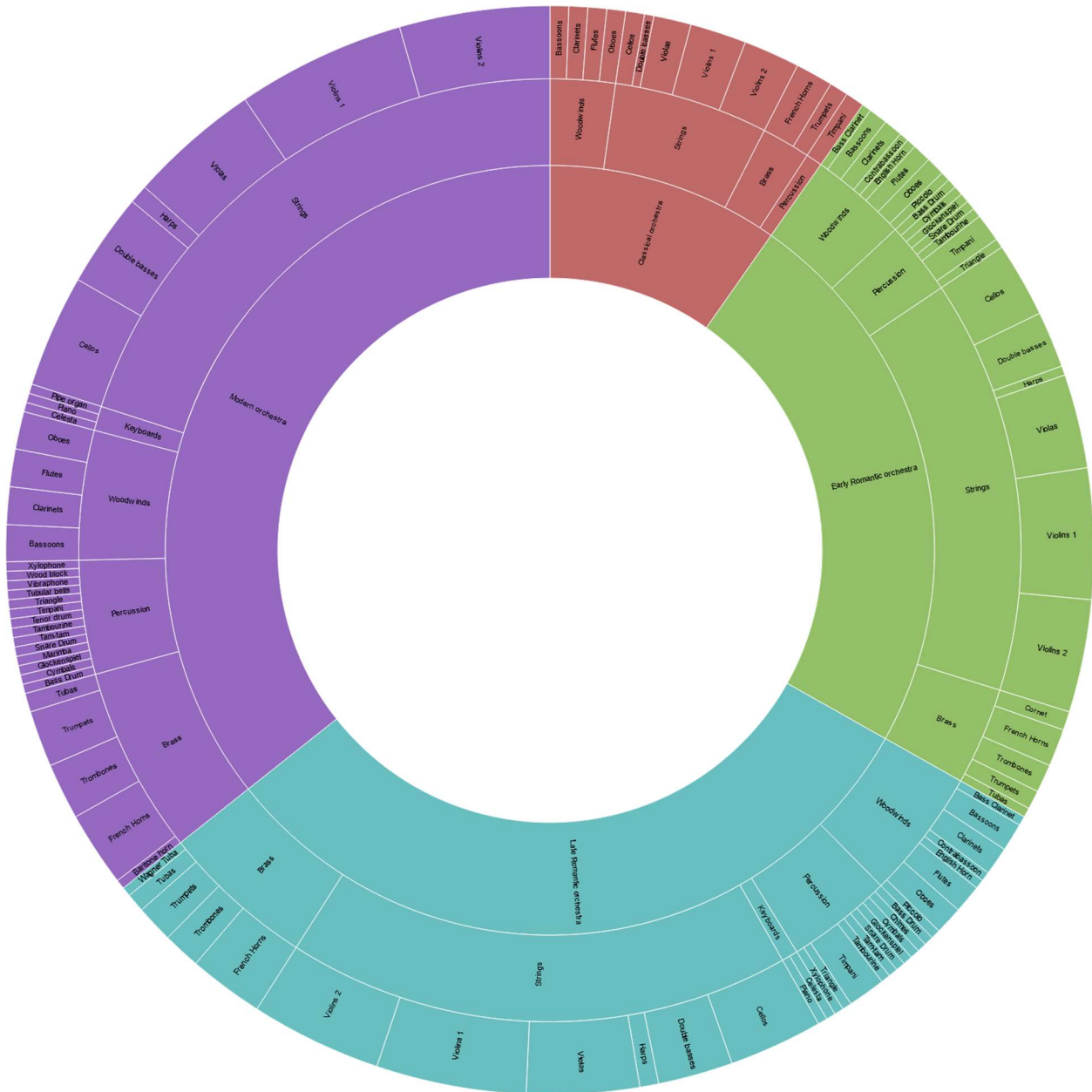
The bar chart is the simplest and most popular out of all the visualisations here. We've seen it a thousand times. Frankly, no report is complete without colourful bar charts!

Bar chart supports four visual patterns as well: X-axis, Height, Groups, Colors.

I mapped Group to the X-axis, Number to Height, Orchestra type to Groups, and Instrument to Colors.

Now here we get to draw similar conclusions as from the Circle map, but here we can actually see the numbers regarding it. In all types of orchestra setting, Strings are the most common type of instruments used, and in that violin, looks to be the most popular.

Sunburst



This visualization supports four visual patterns: Hierarchy, Size, Color, and Label. I mapped the dimensions Orchestra type, Group, and Instrument to Hierarchy. I mapped Number to Size and Orchestra type to Color. This graph provides a better understanding of the Orchestra type. One can see that the Modern Orchestra is the largest subset compared to

the other groups. This visualisation also tells us what instruments fall into what group and what group falls into what orchestra type.

Visualisation Comparison

Graph	Pros	Cons
Treemap	<ul style="list-style-type: none">• Circle Packing, but better• Hierarchy is easily visualised	<ul style="list-style-type: none">• Distortion with more elements• Spacing can make some elements hard to read
Circle Packing	<ul style="list-style-type: none">• Readable• Nested circles allow for meaningful comparison• Strong proportion between elements based on size	<ul style="list-style-type: none">• Messy labelling• Circle sizes can be difficult to differentiate
Bar Chart	<ul style="list-style-type: none">• Most popular charts• Grouping and axes aids comparison	<ul style="list-style-type: none">• Bland• Usable only with discrete data
Sunburst	<ul style="list-style-type: none">• Hierarchy is easily visualised• Readable• Colour aids in distinguishing sections	<ul style="list-style-type: none">• Text can be a bit tight• Might not work well, if it contains too many categories.

Conclusion

When visualising data, we have to ask ourselves, how our visuals aid in the comprehension of our reader about the data. The four graphs do depict the data in ways that should be easily understood by any reader. However, I was disappointed that we're not able to get the legend along with the visualisations from RAW. All these visualisations have their pros and cons, visuals like bar chart have been around for a long time, we all are very used to them by now, but they're not exactly useful to show shares of different attribute values; Treemap, Circle Map, and Sunburst, they all show the shares well, but they all have their own drawbacks. In my opinion, just to depict a concept like market share, I would go with the Sunburst, because it catches the viewers' eyes, the different colours help distinguish between the different sections, and that coloured area gives a good understanding about the shares of any section. Plus, for this project, the drawbacks of Sunburst don't apply because we don't have a lot of distinct attribute values. Therefore, I think that the Sunburst would be a good fit for visualising the share section out of the four graphs we picked. But again, things may change after we learn more about our readers.