



**MONASH** University

**Interactive and dynamic visualization  
of high-dimensional data**

Nicholas Spyris

B.Sc. Statistics, Iowa State University

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2022

Department of Human-Centred Computing

# Contents

|  |      |
|--|------|
| <b>Copyright notice</b>  | v    |
| <b>Abstract</b>  | vii  |
| <b>Declaration</b>   | ix   |
| Publication as part of this thesis . . . . .                                 | ix   |
| Publications and papers during candidature not part of this thesis . . . . . | ix   |
| <b>Acknowledgments</b>   | xi   |
| <b>Preface</b>   | xiii |
| License . . . . .  | xiii |
| <b>1 Introduction</b>  | 1    |
| 1.1 Research questions . . . . .   | 4    |
| 1.2 Methodology . . . . .  | 5    |
| 1.3 Contributions . . . . .  | 6    |
| 1.4 Thesis structure . . . . .   | 8    |
| <b>2 Background</b>  | 9    |
| 2.1 Motivation . . . . .   | 9    |
| 2.2 Multivariate visualization . . . . .                                     | 11   |
| 2.3 Linear projections . . . . .   | 16   |
| 2.4 Evaluating multivariate data visualization . . . . .                     | 18   |
| 2.5 Nonlinear models . . . . .   | 20   |
| 2.6 Local explanations . . . . .   | 21   |
| 2.7 Conclusion . . . . .   | 22   |
| <b>3 A User-Controlled Manual Tour for Animated Linear Projections</b>       | 25   |
| 3.1 Algorithm . . . . .  | 26   |
| 3.2 Oblique cursor movement . . . . .  | 32   |
| 3.3 Package structure . . . . .  | 35   |
| 3.4 Use cases . . . . .  | 37   |
| 3.5 Discussion . . . . .   | 41   |

## CONTENTS

---

|  |            |
|--|------------|
| <b>4 A Study on the Benefit of a User-Controlled Radial Tour for Variable Attribution of Structure</b> | <b>43</b>  |
| 4.1 User study . . . . .   | 45         |
| 4.2 Results . . . . .  | 53         |
| 4.3 Conclusion . . . . .   | 59         |
| <b>5 Exploring Local Explanations of Nonlinear Models Using the Radial tour</b>                        | <b>61</b>  |
| 5.1 SHAP and tree SHAP local explanations . . . . .  | 62         |
| 5.2 The cheem viewer . . . . .   | 65         |
| 5.3 Case studies . . . . .   | 72         |
| 5.4 Discussion . . . . .   | 83         |
| <b>6 Conclusion</b>  | <b>85</b>  |
| 6.1 Contributions . . . . .  | 86         |
| 6.2 Limitations . . . . .  | 87         |
| 6.3 Future work . . . . .  | 88         |
| <b>Bibliography</b>  | <b>93</b>  |
| <b>A Glossary</b>  | <b>107</b> |
| <b>B Supplementary material</b>  | <b>109</b> |
| B.1 Thesis . . . . .   | 109        |
| B.2 Chapter 3, spinifex links . . . . .  | 109        |
| B.3 Chapter 4, user study extended analysis . . . . .  | 110        |
| B.4 Chapter 5, cheem links . . . . .   | 112        |



# **Copyright notice**

© Nicholas Spyrisson (2022).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.



# Abstract

Visualizing data space is a crucial aspect of exploratory data analysis, checking assumptions, and validating model performance. However, visualization quickly becomes complex as the dimensionality of the data or features increases. Traditionally, linear projections have viewed discrete pairs of components to mitigate this complexity. Data visualization *tours* are a class of dynamic linear projections that animate many linear projections over small changes to the projection basis. The permanence of observations between nearby frames potentially conveys more information than discrete orthogonal frames alone.

Tours are categorized by the path of their bases. *Manual tours* uniquely allow for user-controlled steering of a path of bases, where the contributions of individual variables can be changed. The *radial tour* is a specific case of the manual tour, that freezes the angle of movement but allows the magnitude to be varied. This is used in the work reported here. Chapter 3 clarifies the theoretical basis of the radial tour. The details of implementation and illustration of its use are provided. It introduces an open-source **R** package that facilitates creating these tours with a variety of display choices and user controls.

Allowing the analyst to steer a path in the radial tour should enable a better understanding of the variable importance to any structure revealed in a projection. Chapter 4 discusses a within-participant user study comparing the radial tour with current practices: principal component analysis (PCA) and the original type of tour that has no steering. The  $n = 108$  crowdsourced participants performed a variable attribution task describing the separation between clusters. The results find that the radial tour leads to more accurate variable attribution. Participants also reported that the radial tour was their preferred visualization method.

Nonlinear modeling techniques are sometimes referred to as black-box models due to the uninterpretable nature of the model terms. Recent research in Explainable Artificial Intelligence (XAI) tries to bring these models interpretability through *local explanations*. Local explanations are a class of techniques that approximate the linear-variable importance for prediction at one point in the data. Chapter 5 provides a new approach for exploring the variable sensitivity of local explanations using the radial tour. Local explanations can be considered projection bases. The radial tour is used to vary the contribution of a variable to assess its importance for any particular prediction. This is illustrated using four contemporary data examples covering classification and regression. An accompanying R package provides a graphical user interface (GUI) for conducting this analysis.

# Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

## Publication as part of this thesis

The work corresponding to Chapter 3 has been accepted for publication in *The R Journal* ([Spyrison and Cook 2020](#)). While respective studies discussed in Chapters 4 and 5 have not yet been submitted.

I have renumbered and updated sections of the published paper to generate a consistent presentation within the thesis. The code illustrated has been adjusted to highlight the new ggproto API. Introductory and background content has been moved to their respective thesis chapters.

## Publications and papers during candidature not part of this thesis

In addition to the research discussed in the thesis, other notable contributions during my candidature include:

- *The state-of-the-art on tours for dynamic visualization of high-dimensional data* ([Lee et al. 2021](#)). A WIREs Computational Statistics Review of current tour methods. I contributed the writing and visuals discussing manual tours.

## CONTENTS

---

- “*Is IEEE VIS that good?*” *On key factors in the initial assessment of manuscript and venue quality* ([Spyrison, Lee, and Besançon 2021](#)). This alt.VIS workshop paper covers a survey of IEEE VIS authors, how they source articles, decide which to read, and evaluate venue quality. We found low evidence that sentiment changes across academic positions for these topics and provided a discussion on the effects of the “publish or perish” environment, standard author and journal metrics, and difficulty in publishing “null” findings and replication studies.
- *Intraday effect of COVID-19 restrictions on Melbourne electricity consumption* ([Barrow, Chong, and Spyrison 2020](#)). In this Melbourne Datathon entry, we corroborate that the Victorian interday effect on energy consumption did not change but novelly find that the *intraday* distribution of energy consumption does change. There is a statistically significant change in the height between the morning and evening peak energy consumption. We posit this is due to less strict schedules associated with working from home, absence of commute time, and other employment changes. This work was awarded 1st place of hundreds of entries in the insights category of the Melbourne 2020 Datathon.

**Student name:** Nicholas Spyrison

**Date:** 2022-03-03

# Acknowledgments

I would like to express my sincere gratitude to my supervisors, Professor Dianne Cook and Professor Kimbal Marriott, for their support of my Ph.D studies and research, their subject expertise, and their careers supervising and teaching. Thank you for continuously pushing me and my research to new levels. I have enjoyed teaching data visualization, which has been strongly shaped by Di's practical, data-first, visualization-often approach. I will continue to hear Kim's persistent question, "what do we learn from this?" as a reminder not to get lost in the implementation details and regularly step back and analyze if this is the correct object to change. Thanks to the members of the supervisory panel, Associate Professor Bernhard Jenny, Dr. Maxime Cordeil, and Dr. Shirui Pan for their time, engagement, and suggestions throughout this research.

Special thanks to Jieyang Chong and Julie Holden for their proofreading and suggested language clarifications. I thank my fellow Ph.D students, NUMBAT and DVIA lab members, and Pomodoro partners, primarily on the occasion of stimulating discussions and the positive peer pressure of knowing others are around and working. Thanks immensely to those who empathized with me and others, primarily through the hardships of studies and COVID-19. Gratitude to Ying Zhou for her enduring support through the thick of my studies and wavering mental health.

Last but not least, I would like to thank my parents, Doug and Terry, for their continued support and for airing my grievances at odd hours of the day. Thanks to Alan and Claire for their companionship and support now and in our more formative years.

This research was supported in part by an Australian Government Research Training Program (RTP) Scholarship.



# Preface

This thesis has been written using R Markdown with the bookdown package (Xie 2016). All materials required to compile the thesis are available at [github.com/nspyrison/thesis\\_monash\\_phd](https://github.com/nspyrison/thesis_monash_phd). Versions are made available as html and pdf at [nspyrison.github.io/thesis\\_ns/](https://nspyrison.github.io/thesis_ns/) and [github.com/nspyrison/thesis\\_ns/blob/master/docs/thesis\\_ns.pdf](https://github.com/nspyrison/thesis_ns/blob/master/docs/thesis_ns.pdf), respectively.

## License



This code and packages produced art licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).



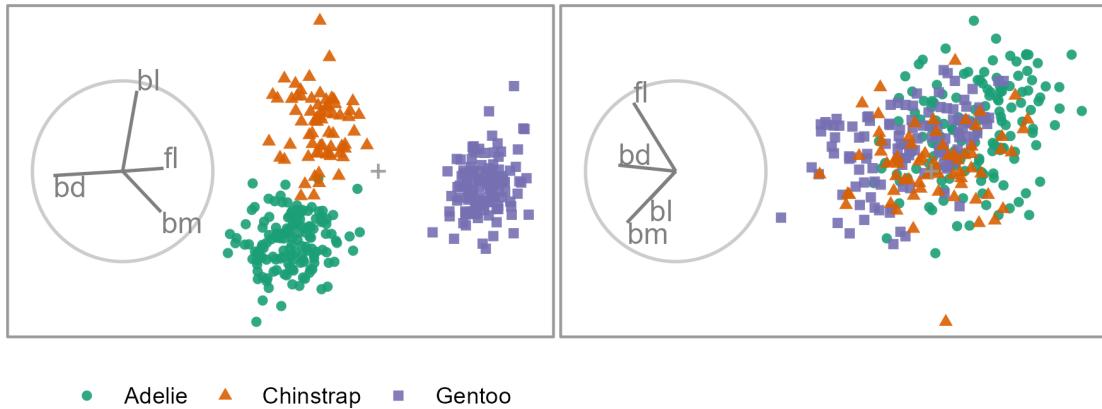
# Chapter 1

## Introduction

Exploratory Data Analysis (EDA) is the process of the initial summarization and visualization of a dataset. EDA is a critical first step that includes checking for realistic values, identifying improper data formats, and revealing insights (Tukey 1977). Wickham and Grolemund (2017) describes the analyst workflow as a series of discrete steps. The data is imported and cleaned before entering into an iterated cycle of transformation, visualization, and modeling. This work focuses on the visualization step, specifically for quantitative multivariate data.

Multivariate data is ubiquitous in contemporary data sets. Multivariate data is found in physics, biology, social sciences, and manufacturing, to name a few (Wang et al. 2018; Huber et al. 2015; Brown 2015; Evans and Boreland 2017, respectively). As the number of variables in the data increases, it becomes increasingly difficult to visualize this space and reveal the structure contained in the data. This thesis addresses the visualization and analysis of multivariate data.

One of the most common and successful approaches to visualize multivariate data is to use linear projection, which approximates  $p$ -dimensional data onto 1-3 display dimensions. In the same way that a 3D object casts a 2D shadow, these projections cast one data orientation onto a lower plane.



**Figure 1.1:** Two linear projections of penguins data. Biplot circles depict the basis indicating the direction and magnitude of the variable contributions. In the left panel, the direction of the separation between the orange and green clusters is in the direction that  $bl$  contributes, meaning that this variable is sensitive to the separation of this cluster. The purple cluster's separation is attributed primarily to  $bd$  and smaller contributions from  $fl$  and  $bm$ . Many other linear orientations do not resolve structures of interest, such as cluster separation (right panel). The Palmer Penguin data ([Gorman, Williams, and Fraser 2014](#); [Horst, Hill, and Gorman 2020](#)) measures four physical variables: bill length ( $bl$ ), bill depth ( $bd$ ), flipper length ( $fl$ ), and body mass ( $bm$ ) for three species of penguins.

Linear projections define new variables called components, which are a linear combination of the original variables. This is essentially an orientation of the direction of the axes. A basis is an orthonormal matrix that maps the variable to the component output space. The basis of a linear projection is frequently illustrated with a biplot ([Gabriel 1971](#)). The biplot shows the magnitude and angle each variable contributes to the resulting display dimensions inscribed in a unit circle, such as in Figure 1.1. Traditional axes-based visuals look at 1- to 3D views of orthogonal variables. In contrast, viewing linear combinations of these variables can reveal features that exist in more than three dimensions.

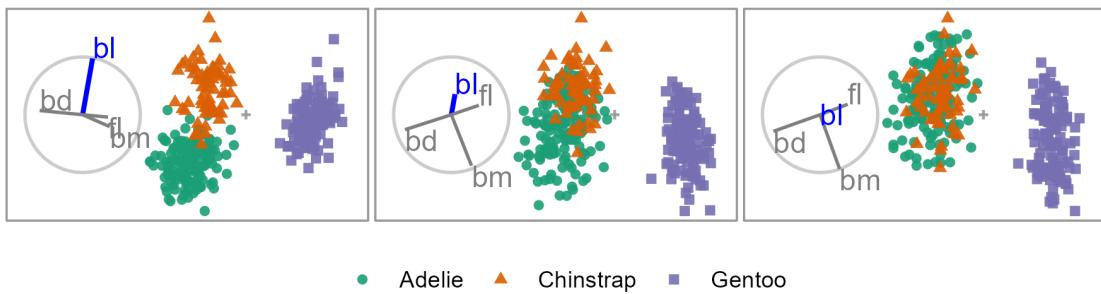
There are many features that an analyst may be interested in when analyzing multivariate data. Shape, spread, clusters separation, outliers, and irregularities are the most common. Not all basis orientations will reveal the features in the data. Thus the choice of basis is essential. Furthermore, the analyst is interested in understanding the variables that reveal features. For instance, Figure 1.1 shows two linear projections of penguin data. The color and shape of the data points distinguish the three penguin species. The linear projection

used in the left panel contains a significant separation of the clusters, while the right panel does not show differences between the species clusters.

Therefore, it makes sense for an analyst to explore many projections with very different bases. However, it can be challenging to link observations across other projections with no relationship to one another. The *tour* is a dynamic visualization that overcomes this difficulty (Cook et al. 2008; Lee et al. 2021). It is a class of linear projections that animate over small changes to the projection basis. A vital feature of the tour is the visual permanence and trackability of the points through the frames. In the shadow analogy, an object such as a barstool will cast a circular shadow if the light is directly above the seat. However, such a shadow does not give the observer sufficient information as it could arise from any number of shapes that contain circular profiles: spheres, cylinders, or circles. However, if the stool was rotated, its legs show in the shadow, giving an intuitive interpretation of the object. Similarly, the rotation of a data object yields information about its structure.

Originally component spaces and the original tour have no means to influence the path of its bases. Cook and Buja (1997) introduced the *manual tour*, offering user control over the basis. By selecting a variable and initializing an additional manipulation dimension onto a projection, the contribution of the variable can be controlled through a rotation of the manipulation space. This user-controlled steering of the basis path enables the analyst to explore what happens to the structure when one variable is removed or the contribution of another is increased.

The manual tour allows the analyst to control the angle and magnitude that a variable contributes to the projection. However, influencing the magnitude is more meaningful as angular manipulations effectively rotate a projection, changing the relative position but not the amounts of the variables. Because of this, this work focuses on a specific manual tour, the *radial tour*, where the manipulation angle is fixed, but a variable's magnitude is changed along the radius. Figure 1.2 shows a radial tour of the penguins data varying the contribution of bill length. The left panel has a total contribution from `b1`, while the middle panel shows half contribution, and bill length has been removed from the right panel. The separation between orange and green clusters was in the direction that `b1`



**Figure 1.2:** A radial tour changing the contribution of bill length ( $bl$ ). When bill length has a considerable contribution, the clusters of orange and green are separated (left and middle). When its contribution is removed, the clusters overlap (right). Because of this, we say that bill length is sensitive to the separation of these two species. An animated version can be viewed at [vimeo.com/676723431](https://vimeo.com/676723431).

contributed and when this contribution is removed, so too is the separation of the clusters. Because of this, this variable is said to be sensitive to the separation of these two clusters.

## 1.1 Research questions

Discerning variable sensitivity to the structure is crucial to understanding which variables contribute to a revealed feature. We conjecture that the user interaction afforded by the radial tour should allow for a more precise exploration of this structure from testing the variable sensitivity to that structure. The over-arching question of interest can, therefore, be stated as:

**Can the radial tour, with user-steering of the basis, help analysts' understanding of the variable sensitivity to structure in the projection?**

While Cook and Buja (1997) sketched the theoretical basis for the manual (and hence radial) tour, some details are missing. Furthermore, there is a lack a publicly available implementation, fully-featured interface design, implementation notes, and have no evaluation of its performance over alternatives.

**RQ 1. How do we define and implement a user interface and interactions for the radial tours to add and remove variables smoothly from 1- and 2D linear data projections?**

At present, the radial tour is not used by analysts. Instead, they would use a single projection to understand the structure, almost always principal components analysis (PCA, [Pearson 1901](#)), which chooses the basis that shows the most variation. Another approach is to use a *grand tour* ([Asimov 1985](#)). The grand tour animates many interpolated frames between randomly selected target bases. Neither PCA nor the grand tour provides a means for manually manipulating a desired variable's contribution to the basis. We wish to investigate if the basis-steering of the radial tour facilitates a better understanding of variable sensitivity to the structure.

**RQ 2. Does the use of the interactive radial tour improve analysts' understanding of the relationship between variables and structure in 2D linear projections compared to existing approaches?**

Complex nonlinear models are also being applied more frequently to predict or classify with many predictors. While these models lead to increased accuracy over linear models, they suffer from a loss of the interpretability of their variables. One aspect of eXplainable Artificial Intelligence (XAI, [Adadi and Berrada 2018](#); [Arrieta et al. 2020](#)) tries to preserve the interpretability of such models through local explanations. These explanations are essentially linear variable importance in the vicinity of one observation of a model. That is, the extent that variables help the model explain the difference between the observed means and an observation's prediction. The user control from the radial tour potentially allows an analyst to better understand the model and the support of these local explanations.

**RQ 3. Can the radial tours be used in conjunction with local explanations to improve the interpretability of black-box models?**

## 1.2 Methodology

The research corresponding to RQ 1 entails *algorithm & software design* ([Kleinberg and Tardos 2006](#)), and adapts the algorithm from [Cook and Buja \(1997\)](#).

To address RQ 2, *experimental design* ([Winer 1962](#)) is used. A task and measure must be defined that is suitable to evaluate the radial tour against alternatives. The experimental factors and their levels must be selected and randomly assigned as a randomized controlled

---

trial to measure the efficacy of user-controlled radial tours compared with two benchmark methods.

The research responding to RQ 3 involves *design science* (Hevner et al. 2004). It is not obvious how to combine a radial tour with a nonlinear model. A local explanation approximates the linear variable importance in the vicinity of one observation. Two novel interactive visualizations need to be developed. The first should visually facilitate the selection of observations to explore. The second should extend the biplot to show the distribution of local explanations from all observations while a radial tour is created starting from the attribution of one observation. This will examine the variable sensitivity to the structure identified in the local explanation.

## 1.3 Contributions

The contributions resulting from the research to address these research questions can be split into scientific knowledge and software contributions:

### 1.3.1 Scientific knowledge

- Radial tour algorithm
  - Refined and clarified the steps to producing a radial tour based on the use of the Rodrigues' rotation formula extends the approach from [Rodrigues \(1840\)](#).
  - The algorithm makes the radial tour in the animation and interactive systems simpler to implement.
  - Provides new examples of usage.
- A user study comparing the radial tour's efficacy against two alternatives — PCA and the grand tour, the first empirical evaluation of the radial tour.
  - Creation of supervised classification task to assess the variable attribution to the separation of two clusters.
  - As tested over experimental factors: location, shape, and dimensionality.
  - Definition of an accuracy measure to evaluate this task.

- Results: strong evidence that the radial tour increases the accuracy of this task by a sizable amount and minor evidence to suggest a moderate increase in accuracy of the grand tour over PCA.
- Mixed model regression helps to attribution the source of the error accounting for the variability of participant’s skills and the difficulty of the simulation by chance.
- A novel interactive visualization using the radial tour to explore the variable sensitivity of local explanations from nonlinear models, part of XAI.
  - A global view approximates the variable space, attribution space, and residual plot side-by-side, which serves to identify observations of interest.
  - Selected observation’s normalized variable attribution is used as a projection basis.
  - Explore the support of the local explanation using the radial tour; the variable sensitivity to the structure identified tests the range of contributions supporting the explanation.

### 1.3.2 Software

- **spinifex**, an **R** package for transforming data, performing manual tours, and extending any tour’s display and animation exportation.
  - Facilitates the transformation of numeric variables in the data.
  - Identify various bases finding various features of the data.
  - Creation of manual tours allows analyst steering of the basis to explore the variable sensitivity to the structure.
  - Layered composition of tour displays that mirrors the approach in **ggplot2** (Wickham 2016), interoperable with tours made by **tourr** (Wickham et al. 2011).
  - Exporting rendered animation either to interactive **html** widgets with **plotly** (Sievert 2020) or to **gif**, **mp4**, and other video formats with **ganimate** (Pedersen and Robinson 2020).
  - Interactive **shiny** (W. Chang et al. 2021) application to preprocess data and explore. Users can choose from six supplied datasets or upload their own.

- Vignettes and code examples help users get up to speed.
- **cheem**, an **R** package that facilitates the exploration of local explanations of nonlinear models through the radial tour.
  - Preprocessing: given a tree-based model, calculate the tree SHAP local explanation (via **treeshap**, Kominsarczyk et al. 2021) of all observations, and calculate statistics that accent the separability of this space.
  - Visualization of approximations of the data space, attribution space, and residual plot side-by-side with linked brushing, hover tooltips, and tabular display facilitates the selection of observations to explore.
  - Use of the radial tour changes variable contribution to test the support of the variable contribution in agreement with the explanation.
  - Interactive application facilitates this analysis for several prepared datasets or user preprocessed data.
  - A vignette and code examples help users get up to speed.

## 1.4 Thesis structure

The remainder of the thesis is organized as follows: Chapter 2 covers various visualization techniques before introducing related studies, nonlinear models, and their interpretability issues. Chapter 3 discusses the theory and implementation of the radial tours in the package **spinifex**. Chapter 4 discusses a user study evaluating the radial tour’s efficacy compared with PCA and the grand tour. Chapter 5 discusses the **cheem** package, which extends radial tours to improve the interpretability of nonlinear models. Lastly, Chapter 6 concludes with some takeaways and a discussion of limitations and possible extensions.

# Chapter 2

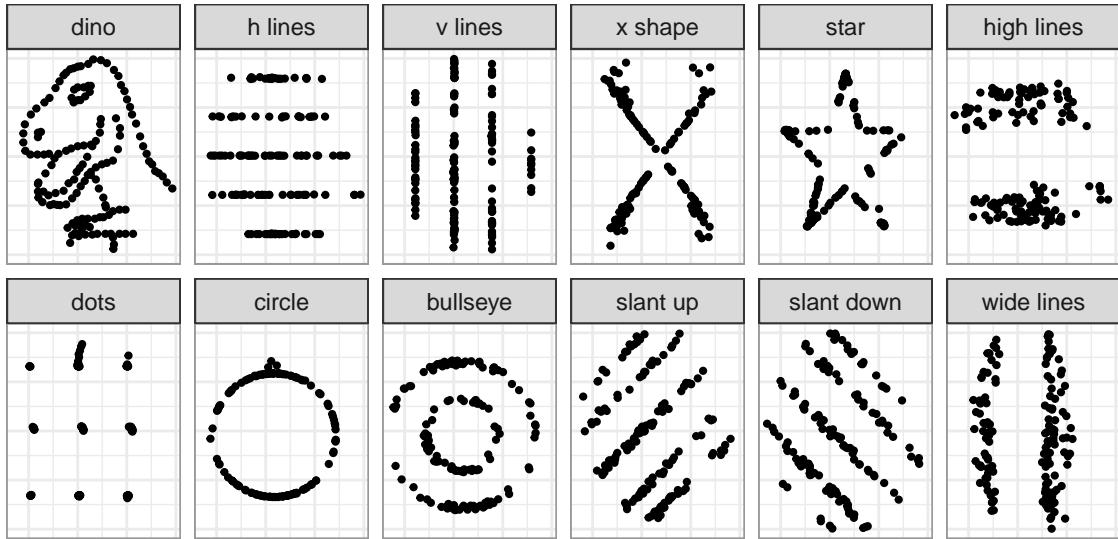
## Background

The last chapter discussed the importance of data visualization despite the complexity of viewing high-dimensional data and outlined the research questions to be addressed around the user-control of the radial tour. This chapter first motivates data visualization in general and the importance of interaction. It continues on illustrating common visualizations for quantitative multivariate data before turning to dimension reduction, including further discussion of tours. Then empirical evaluations of multivariate visuals are covered. The chapter concludes with nonlinear models and the extension of their interpretation with local explanations.

### 2.1 Motivation

Visualization is much more robust than numerical summarization alone ([Anscombe 1973](#); [Matejka and Fitzmaurice 2017](#)). Figure 2.1 illustrates this. The data sets have a quite different structure revealed in their visualization but not in the summary statistics. All data sets have the same means, standard deviations, and correlation.

Interaction is an essential aspect of modern data visualization ([Batch et al. 2019](#); [Card, Moran, and Newell 1983](#); [Marriott et al. 2018](#)). Interaction facilitates accessing increasing amounts of data and amplifies cognition through control and input ([Dimara and Perin 2019](#)). [Munzner \(2014\)](#) posits that responsive and fast computer graphics allow us to move



**Figure 2.1:** *The datasaurus dozen is a modern data set illustrating Anscombe’s point that summary statistics cannot always adequately summarize data content. Patterns in the scatterplots are evident despite the data having the same summary statistics. Unless the data is plotted, one would never know that there were such radical differences.*

beyond paper and static resources. Interaction is key to navigating within views of sizable data and linking observations and features across these views.

We focus on multivariate data interactions with coordinated views, linked brushing, and tooltip display. Coordinated and multiple views (Roberts 2007) (also known as ensemble graphics, Unwin and Valero-Mora 2018) use several types of visuals that give a more comprehensive understanding than any one visual portrays in isolation. The linking of observations between different views and animation frames is facilitated by linked brushing (Becker and Cleveland 1987). In linked brushing, selected observations in one view are colored across all views, allowing these selections to be tracked and correlated across other views and frames. Linked brushing has proved to be helpful in animated tours (Arms, Cook, and Cruz-Neira 1999; Laa, Cook, and Valencia 2020; Lee, Laa, and Cook 2020). Tooltip displays upon the cursor hovering over observation can display identification information and other associated values to aid with point identification and more detailed information to be accessed. Interactive selection of parameters extends the breadth of analysis, and animation interactions such as play, pause, and frame selection are regularly used.

## 2.2 Multivariate visualization

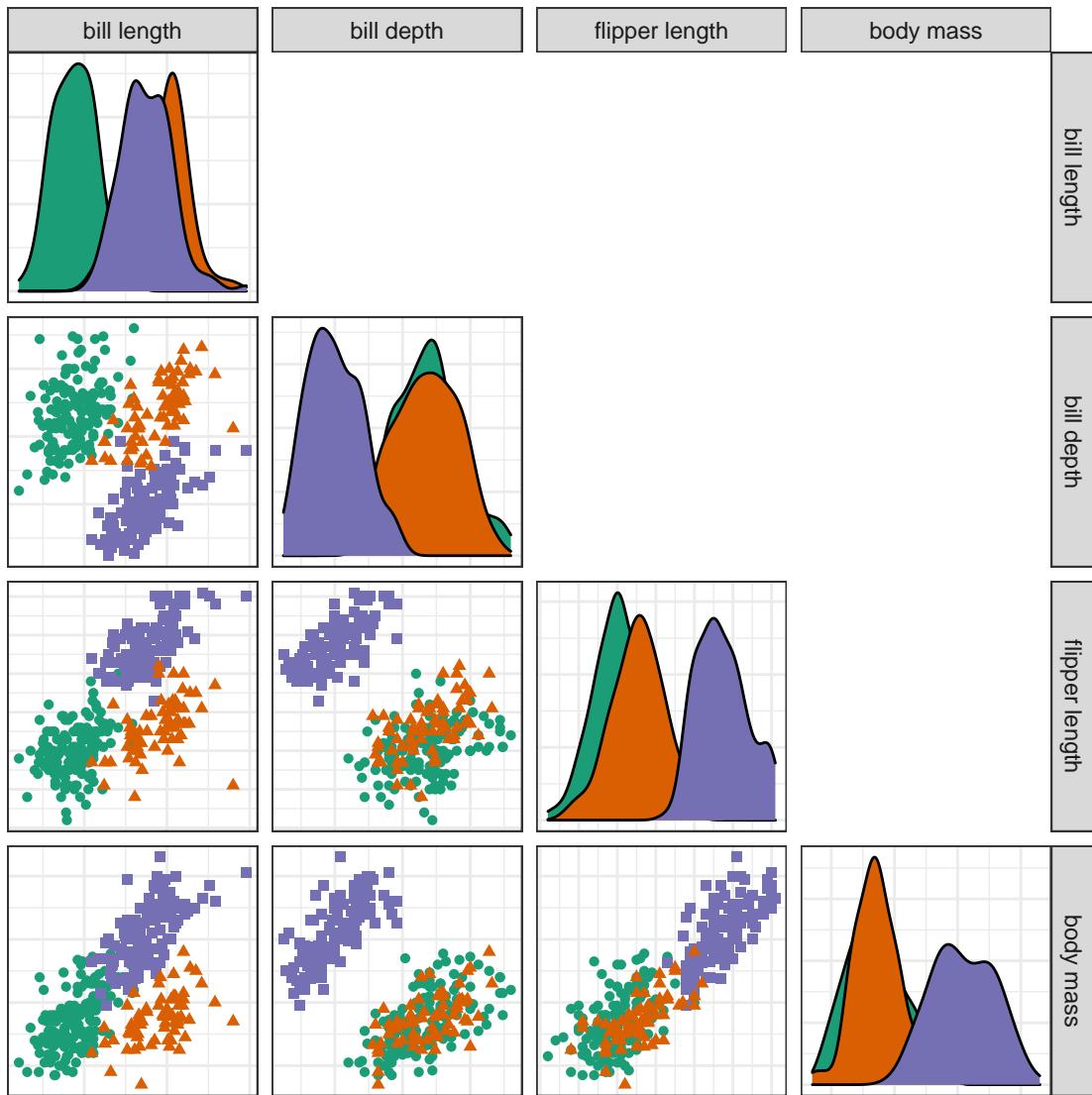
In this thesis, we are concerned with the visualization of multivariate data. Specifically, we are interested in quantitative multivariate data. We assume that our data consists of  $n$  observations of  $p$  variables. Generally  $n > p$  with many more observations than variables. While written as though operating on the original variables, the discussion below could also be applied to reduced component spaces (such as PCA approximation in a few components) or feature decomposition of data not fitting this format. [Kang, Hyndman, and Smith-Miles \(2017\)](#) provide an excellent example of decomposing time series data in a quantitative feature matrix.

[Grinstein, Trutschl, and Cvek \(2002\)](#) illustrate many multivariate visualization methods. In particular, this work shows examples of actual visuals. [Liu et al. \(2017\)](#) give a good classification and taxonomy of such methods. The content below focuses on the most common and the most relevant visuals. Illustrations are provided with the Palmer penguins data that was used in the Introduction. This data contains 333 observations of three penguin species across four physical measurements: bill length, bill depth, flipper length, and body mass. Observations were collected between 2007 and 2009 near Palmer Station, Antarctica. The penguins data is a good substitute for the over-used iris data.

### 2.2.1 Scatterplot matrices

An analyst could look at  $p$ -univariate histograms or density curves. Extending this idea, pairs of variables can be exhaustively viewed. Combining these brings us to scatterplot matrices, also known as SPLOM ([Chambers et al. 1983](#)). In a scatterplot matrix, variables are arranged across the columns and rows. The diagonal elements show univariate densities, while off-diagonal positions show scatterplot pairs, as in Figure 2.2. This is useful for getting a handle on the range of the variables but is not going to scale well when the number of variables  $p$  is large. Such visualization will only partially resolve features that could be better show information from more dimensions.

As  $n$  increases, scatterplot displays also suffer from occlusion as the points overlay each other. This is typically addressed in a few ways. One method is to decrease points' opacity,



**Figure 2.2:** A scatterplot matrix shows univariate densities and all pairs of bivariate scatterplots. The panels show partial cluster separation, indicating that these variables contain discerning information. This approach is suitable for quickly exploring the range of the data but will not reveal features in more than two dimensions. It is a good exploratory visual but does not scale well with increasing variables.

allowing more layers to be seen. Another approach is to change the geometric display, such as 2D density contour or an aggregated heatmap (illustrated in Figure 2.5). Aggregated displays typically render faster and scale better with increasing observations. Or, if needed, visualization can be performed on a representative subset of the data.

### 2.2.2 Parallel coordinate plots

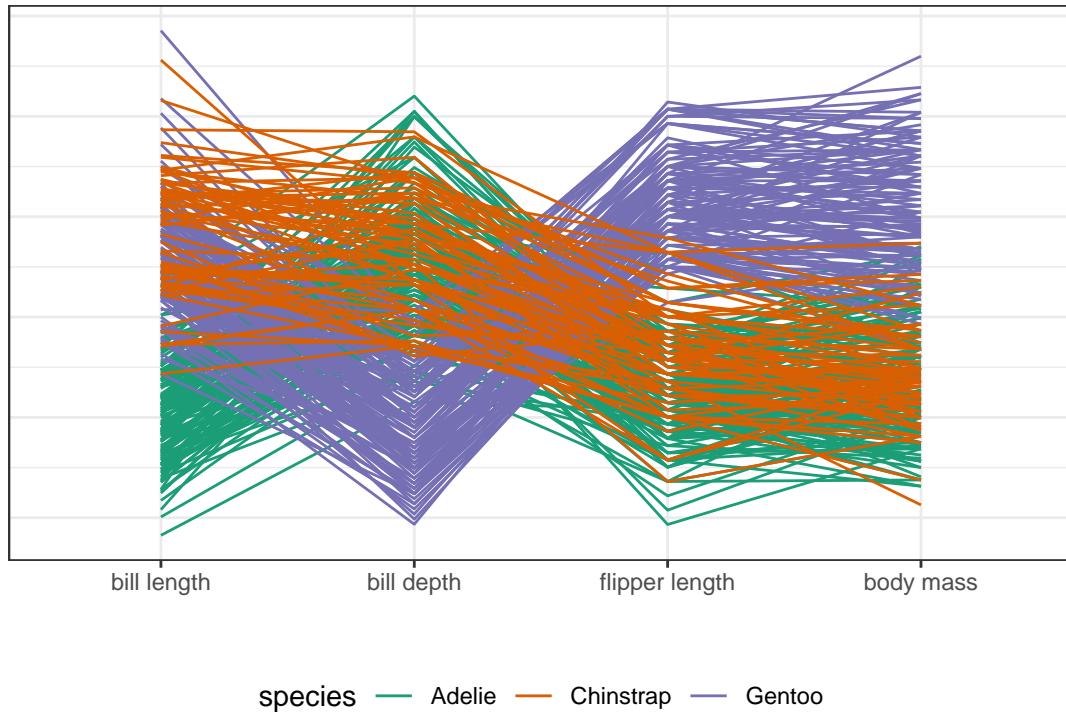
In scatterplot matrices, each observation repeated across all panels. In contrast, observation-linked visuals have a single line or glyph for each observation. In parallel coordinate plots ([Ocagne 1885](#)), variables are arranged horizontally, and lines connect observations after being transformed to a common scale such as quantiles or z-score (standard deviations away from the mean). Figure 2.3 illustrates this method.

Parallel coordinate plots scale much better with dimensions than scatterplot matrices but more poorly with observations. They also suffer from an asymmetry with the variable order. That is, changing the order of the variables may lead to very different conclusions. The  $x$ -axis is also used to display variables rather than the values of the observations. This restricts the amount of information that can be interpreted between variables. [Munzner \(2014\)](#) asserts that position is the more human-perceptible channel for encoding information; we should prefer to reserve it for distinguishing between values of the observations rather than the arrangement of variables.

The same issues persist across displays that map observations into  $n$  glyphs or pixel heatmaps. Examples of these include star plots ([Chambers et al. 1983](#)), pixel-based visuals ([Keim 2000](#)), and Chernoff faces ([Chernoff 1973](#)). Like parallel coordinate plots, these other visuals scale quite poorly with increasing observations. However, because these visuals scale well with the number of variables, they may be candidate visualizations for low  $n$ , high  $p$  data.

### 2.2.3 Dimension reduction

The other main approach for visualizing quantitative multivariate data is dimension reduction. This involves a function mapping  $p$ -space onto a lower  $d$ -dimensional space. Dimension reduction is separated into two categories, linear and nonlinear. The linear case



**Figure 2.3:** Parallel coordinate plots put variables on a common scale and position them side-by-side with lines connecting observations. Some variation between the clusters can be seen corroborating their importance to explaining cluster separation. This approach scales relatively well with the number of variables but poorly with the number of observations.

spans all affine mathematical transformations, essentially any function where parallel lines stay parallel. Nonlinear transformations complement the linear case, think transformations containing exponents or interacting terms.

Examples in low dimensions are relatable. For instance, shadows are linear projections of a 3-dimensional object down to a 2D shadow. Linear perspective drawings are another instance. An example of a nonlinear transformation is that of 2D maps of the globe. A common example is the Mercator projection, a rectangular display where the area is proportionally distorted with the distance away from the equator (Snyder 1987). Other distortions are created when the surface is unwrapped into an elongated ellipse. Yet others create non-continuous gaps on land or oceans to minimize the distortion of targeted areas. Snyder lists over 200 different projections that distort the surface to display as a map, each with unique properties distorting the 3D surface. However, despite familiarity with map

projections, users find it difficult to understand the distortions they introduce (Hennerdal 2015).

As illustrated by map projections, nonlinear projections can be challenging to understand. Various computational quality metrics, such as Trustworthiness, Continuity, Normalized stress, and Average local error, have been introduced to describe the distortion of the space (Espadoto et al. 2021; Gracia et al. 2016; Maaten, Postma, and Van den Herik 2009; Venna and Kaski 2006). To quote Panagiotelis (2020), “All nonlinear projections are wrong, but some are useful,” a play on George Box’s quote about models (“All models are wrong, but some are useful,” Box 1976). The distortions are hard to interpret and, thus, communicate.

Furthermore, nonlinear projections have hyperparameters (absent from linear methods) that control how the spaces are distorted to fit into fewer dimensions. These introduce a degree of subjectivity into the resulting projection. Opinions differ on how to best deal with hyperparameters. Probst, Wright, and Boulesteix (2019) discuss tuning strategies. Others compare implementation default values (Gijsbers et al. 2021; Pfisterer et al. 2021). Probst, Boulesteix, and Bischl (2019) look at the sensitivity of hyperparameters to the performance of a model. Automated machine learning takes a programmatic approach to hyperparameter tuning (Feurer et al. 2015; Hutter, Kotthoff, and Vanschoren 2019; Yao et al. 2019). Due to the difficulty of interpreting nonlinear mappings and the added subjectivity of hyperparameter selection, this thesis focuses on linear visualization techniques.

### 2.2.4 Intrinsic data dimensionality

One way dimension reduction is used is to project multivariate data onto 1-, 2-, or 3D space and visualize the results. However, it can also be used as a preprocessing step for analysis in  $d$ -dimensions. The intrinsic dimensionality of data is the number of variables needed to minimally represent the data (Grinstein, Trutschl, and Cvek 2002). Intrinsic data dimensionality is an essential consideration of dimension reduction. Consider a Psychology survey consisting of 100 questions about the Big Five personality traits. The data consists of 100 response variables, while the theory would suggest the intrinsic dimensionality is five. It thus makes sense to project onto five dimensions and analyze and visualize this embedded

space rather than the original space as this mitigates the exponentially increasing volume of the view space and the computational load with minimal information loss.

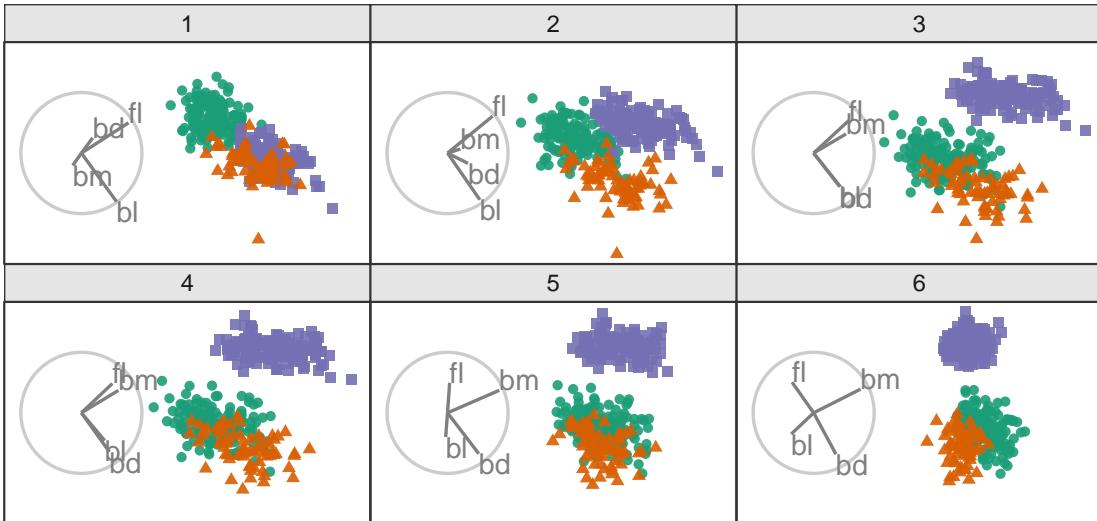
## 2.3 Linear projections

Linear projections map a higher  $p$ -dimensional space onto a smaller  $d$ -space with an affine mapping (where parallel lines stay parallel). A projection is the resulting space of the data multiplied by a *basis*  $Y_{n \times d} = X_{n \times p} \cdot A_{p \times d}$ . This basis is an orthonormal matrix that explains the orientation and magnitudes that variables contribute to the resulting space. This basis is often illustrated as a biplot ([Gabriel 1971](#)), where variable contributions are inscribed in a unit circle showing the direction and the magnitude of contribution as presented in the previous chapter and in Figure 2.4.

A standard linear projection is principal component analysis (PCA, [Pearson 1901](#)), which creates a component space ordered by descending variation. It uses eigenvalue decomposition to identify the basis. These components are typically viewed as discrete orthogonal pairs, commonly approximated as several components. The exact number of components to keep is subjective but typically guided by a screeplot ([Cattell 1966](#)). Scree plots illustrate the decreasing variation contained in subsequent components. The analyst then identifies an “elbow” in this plot. PCA is also commonly used in preprocessing to reduce the number of dimensions to embed the data in a space corresponding to the intrinsic dimensionality.

### 2.3.1 Tours, animated linear projections

In a static linear projection, there is only one basis. However, a single projection may not reveal the structure of interest. In contrast to this, a data visualization tour shows multiple projections by animating between small changes in the basis. In the shadow analogy, structural information of an object is gained by watching its shadow change due to its rotation. An analyst similarly gains structural information by watching a continuous projection over changes to the basis (essentially a rotation of the data). There are various types of tours that are classified by the generation of their basis paths. We enumerate a few related to this work. A more comprehensive discussion and review of tours can be found in the works of [Cook et al. \(2008\)](#) and [Lee et al. \(2021\)](#).

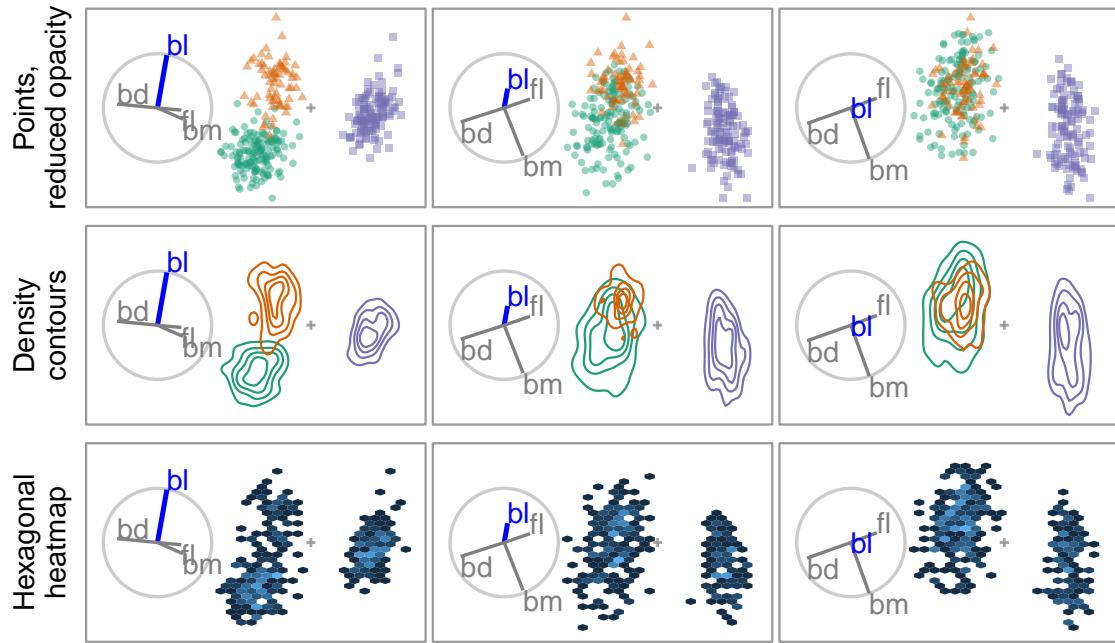


**Figure 2.4:** Frames from a *grand tour*. Biplots (grey circles) illustrate the direction and magnitude that variables contribute. In the *grand tour*, target bases are selected randomly. Tours animate linear projections over small changes in the basis. The observation permanence between frames is an essential distinction in tours. An animation can be viewed at [vimeo.com/676723441](https://vimeo.com/676723441).

Regardless of the type of tour, they generate a sequence of target bases, and then the tour must interpolate intermediate bases between consecutive bases. This interpolation is performed along a geodesic path between the bases. Geodesic refers to the shortest path (on a  $p$ -sphere of possible bases). This ends up being slightly curved in 2D representation for the same reason that flight paths appear curved on maps. The interpolation of bases at small enough angles is foundational for the trackability of observations between frames.

Originally in a *grand tour* (Asimov 1985), several target bases are randomly selected. Figure 2.4 illustrates six frames of a *grand tour*. The *grand tour* is suitable for exploratory data analysis in that it will show bases with widely varying contributions but lacks a means of steering the tour and controlling the choice of bases.

In contrast, the *manual tour* (Cook and Buja 1997) allows the analyst to change the contribution of a selected variable. It does so by initializing a manipulation dimension on a 1- or 2D projection which can then be rotated to alter the contribution of a selected variable. This work focuses on the *radial tour* sub-variant, where the contribution angle is fixed, and the magnitude of contribution along the radius is controlled. We saw an example of this in Figure 1.2.



**Figure 2.5:** *Radial tour across three geometric displays. Changing the point opacity or geometric display to density contours or aggregated heatmap are common ways to mitigate occlusion caused by dense observations. The radial tour removes the contribution of *bl* and, with it, the separation between the orange and green clusters. Heatmap display is not the best choice to show supervised cluster separation but can be helpful to see structure in dense data.*

Figure 2.5 illustrates the same radial tour shown in the previous chapter across three geometric displays designed to overcome occlusion issues when there are a large number of observations. The use of density contours and aggregated heatmap displays help to mitigate the occlusion of dense observations and is compatible with any scatterplot display.

## 2.4 Evaluating multivariate data visualization

Definitions and surveys of quality metrics for the distortion of nonlinear reduction are given in [Bertini, Tatu, and Keim \(2011\)](#). The latest, most comprehensive quantitative surveys are discussed in [Espadoto et al. \(2021\)](#) and [Nonato and Aupetit \(2018\)](#). The former compares 44 dimension reduction techniques across 18 datasets over seven metrics. While the former papers mention tours, they are absent from quantitative surveys.

Some studies compare visualizations across complete contributions of variables. [C. Chang, Dwyer, and Marriott \(2018\)](#) conduct an  $n = 51$  participant study comparing parallel

coordinate plots and SPLOM either in isolation, sequentially, or as a coordinated view. Accuracy, completion time, and eye focus were measured for six tasks. Three tasks were more accurate with SPLOM and three with parallel coordinates, while the coordinated view was usually marginally more accurate than the max of the separate visuals. Cao et al. (2018) compare unstandardized line-glyph and star-glyphs with standardized variants (with and without curve fill). Each of the  $n = 18$  participants performs 72 trials across the six visuals, two levels of dimensions, and two levels of observations. Visuals with variable standardization outperform the unstandardized variants and the radial star-glyph reportedly outperformed line-variant.

Other studies have investigated the relative benefits of projecting to 2- or 3D scatterplots in PCA-reduced spaces. Gracia et al. (2016) conducted an  $n = 40$  user study comparing 2- and 3D scatterplots on traditional 2D monitors. Participants perform point classification, distance perception, and outlier identification tasks. The results are mixed and primarily have small differences. There is some evidence to suggest a lower error in distance perception from a 3D scatterplot. Wagner Filho et al. (2018) performed an  $n = 30$  within participants study on PCA reduced space using scatterplot displays between 2D on monitors, 3D on monitors, and 3D display with a head-mounted display. None of the tasks on any dataset lead to a significant difference in accuracy. However, the immersive display reduced effort and navigation, resulting in higher perceived accuracy and engagement.

Some studies use an expert or cohort coding. Sedlmair, Munzner, and Tory (2013) instead use two expert coders to evaluate 75 datasets and four dimension reduction techniques across 2D scatterplots, 2D scatterplot matrices, and interactive 3D scatterplots. They suggest a tiered guidance approach finding that 2D scatterplots are often sufficient to resolve a feature. If not, try an alternative dimension reduction technique before going to scatterplot matrix display or concluding a true negative. They find that interactive 3D scatterplots help in relatively rare cases. Lewis, Van der Maaten, and Sa (2012) compare across three cohorts: experts, uninformed novices, and informed novices ( $n = 5 + 15 + 16 = 36$ ). Participants were asked their opinion of the quality of nine different embedding for each of several data sets. Expert opinion is reportedly more consistent than the novice groups, though this is confounded with the different sample sizes. Interestingly, cohort responses correlated with

---

different quality metrics. Positive ratings from the expert group correlated strongest with the Trustworthiness metric.

Tours are absent from studies calculable quality measures. However, [Nelson, Cook, and Cruz-Neira \(1998\)](#) compare scatterplots of grand tours on a 2D monitor with 3D display (stereoscopic, not head-mounted) over  $n = 15$  participants. Participants perform clusters detection, dimensionality, and radial sparseness tasks on six-dimensional data. They find that stereoscopic 3D leads to more accuracy for cluster identification, though interaction time greatly increased in the 3D case.

## 2.5 Nonlinear models

So far, the thesis has focused on exploratory data analysis. Another core part of data analysis is fitting *regression* models of quantitative data ([Galton 1886](#)) and *classification* models for discrete categories ([Fisher 1935](#)). There are different reasons and emphases when considering to fit a model. [Breiman \(2001b\)](#), reiterated by [Shmueli \(2010\)](#), taxonomize models based on their purpose. *Explanatory* modeling is done for some inferential purpose, while *predictive* modeling focuses more narrowly on the performance of some objective function. The intended use has important implications for model selection and development. In explanatory modeling, interpretability is vital for drawing inferential conclusions. Nonlinear models range from additive models with at least one polynomial term to more complex machine learning models such as random forest, gradient boosting, or neural network to name a few ([Breiman 2001a; Friedman 2002; Anderson 1995](#), receptively).

Nonlinear models have many or complex interactive terms, which cause an opaqueness to the interpretation of the variables. This difficulty in interpreting the terms in complex nonlinear models sometimes leads to them being referred to as black box models. However, the use and prevalence of black box models is not without controversy ([O’Neil 2016; Kodiyan 2019](#)). And the loss of interpretation presents a challenge.

Interpretability is vital for exploring and protecting against potential biases in any model (e.g. sex – [Dastin \(2018\); Duffy \(2019\)](#), race – [Larson et al. \(2016\)](#), and age – [Díaz et al. \(2018\)](#)). For instance, models regularly pick up on biases in the training data where such

classes correlate with changes in the response variable. This bias is then built into the model. Variable-level (feature-level) interpretability of models is essential in evaluating and addressing such biases.

Another concern is data drift, where a shift in the range of the explanatory variables (features or predictors). Some nonlinear models are sensitive to this and do not extrapolate well outside the support of the training data. Maintaining variable interpretability is also essential to address issues arising from data drift.

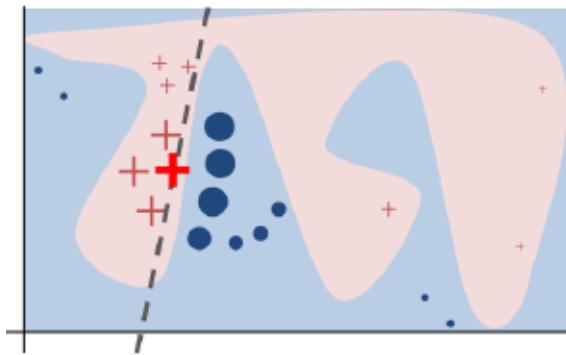
## 2.6 Local explanations

Explainable Artificial Intelligence (XAI) is an emerging field of research that aims to increase the interpretability of nonlinear models. A common approach is *local explanations*, which attempt to approximate linear variable importance in the vicinity of one observation (instance). This is a linear measure indicating which variables are essential for distinguishing between the mean of the data and the prediction near one observation. Because these are point-specific, the challenge is to comprehensively visualize them to understand a model.

Consider a highly nonlinear model. It can be hard to determine which variables in the data are sensitive to changes in the classification or sizable changes in a residual. Local explanations shed light on these situations by approximating linear variable importance in the vicinity of a single observation. Figure 2.6 motivates local explanations where the analyst wants to know the variable attribution for a particular observation close to the classification boundary in a nonlinear model.

A comprehensive summary of the taxonomy and literature of explanation techniques is provided in Figure 6 of Arrieta et al. (2020). It includes a large number of model-specific explanations such as deepLIFT (Shrikumar et al. 2016; Shrikumar, Greenside, and Kundaje 2017), a popular recursive method for estimating importance in neural networks. There are fewer model-agnostic explanations, of which LIME, (Ribeiro, Singh, and Guestrin 2016) SHAP, (Lundberg and Lee 2017), and their variants are popular.

These observation-level explanations are used in various ways depending on the context. In image classification, a saliency map indicate necessary pixels for the resulting classification



**Figure 2.6:** Illustration of a nonlinear classification model. An analyst may want to know the variable importance in the vicinity of the highlighted red cross. Understanding this attribution elucidates how the variables influence this point that is precariously close to the classification boundary. Local explanations approximate this linear attribution in the vicinity of one observation. Figure from Ribeiro, Singh, and Guestrin (2016).

(Simonyan, Vedaldi, and Zisserman 2014). For example, snow may be highlighted when distinguishing if a picture contains a wolf or husky (Besse et al. 2019). In text analysis, word-level contextual sentiment analysis can be used to highlight the sentiment and magnitude of influential words (Vanni et al. 2018). In the case of numeric regression, they are used to explain variable additive contributions from the observed mean to the observation’s prediction (Ribeiro, Singh, and Guestrin 2016).

## 2.7 Conclusion

This chapter discussed the motivation for data visualization showing more information than statistic summarization alone. User interaction is important for linking observation and auxiliary information between coordinated views and animation frames. Several visuals were discussed before turning to dimension reduction. Because of the opaqueness of the distortions and added subjectivity from nonlinear dimension reduction, this thesis focused more on viewing many linear projections as the basis changes in tours. The radial tour with user-control steering is of particular interest.

Metric and empirical evaluations of multivariate data were discussed. Nonlinear models, and XAI’s use of local explanations to extend the interpretability of models were also covered. There is an absence of studies comparing animated tours with alternative visualizations.

The following chapters respectively address the three research questions covered in the Introduction. Chapter 3 discusses the implementation of a package that facilitates the creation of radial tours and extends the display and exporting of tours in general. Chapter 4 covers the first user study evaluating the radial tour compared with two common alternatives. Chapter 5 introduces a novel analysis that explores local explanations of nonlinear models with the radial tour.



## Chapter 3

# A User-Controlled Manual Tour for Animated Linear Projections

Dynamic low-dimensional linear projections of multivariate data known as *tours* are essential for exploring multivariate data and models. The **R** package **tourr** facilitates several types of tours: grand, guided, little, local, and frozen. Each of these can be viewed in a development environment, or their basis path can be saved for later consumption. This chapter describes a new package, **spinifex**, which creates *manual tours* (Cook et al. 1995) of multivariate data. In a manual tour, an analyst controls the contribution of a variable to the projection. Controlled manipulation is important to explore a variable's sensitivity to the structure of an identified feature. The *radial tour* is a sub-variant that fixes the contribution angle and alters the magnitude along the radius. The radial tour is applied to particle physics data to illustrate the structure's sensitivity in a projection to specific variable contributions. Additionally, this package creates a ggproto API for composing any tour that mirrors the layered additive approach of **ggplot2**. Tours can then be animated and exported to various formats with **plotly** or **ganimate**.

The chapter is organized as follows. Section 3.1 provides an illustrated detailed description of a 2D manual tour and fills in the previously absent details to solve the 3D rotation matrix used in 2D manual tours. This proves a scaffolding for the extension to solving a 4D rotation matrix for a 3D manual tour. Section 3.2 includes compact Algorithms

for capturing oblique movements from a cursor. Section 3.3 discusses the functions and code usage to perform radial tours and compose a layered display of tours. Section 3.4 illustrates a use case of the radial tour facilitating sensitivity analysis on high-energy physics experiments (Wang et al. 2018). Section 3.5 concluded with a discussion of this chapter.

## 3.1 Algorithm

The types of manipulations of the manual tour can be thought of in several ways:

- *radial*: fix the direction of contribution, and allow the magnitude to change.
- *angular*: fix the magnitude, and allow the angle or direction of the contribution to vary.
- *horizontal, vertical*: changing the contribution in horizontal or vertical directions.
- *oblique*: paths deviating from these movements, such as being captured from the movement of a cursor.

Angular manipulations are homomorphic in that they show the same information while rotating the frame. More interesting is a change in the magnitude of the contribution, changing the radius along the original contribution angle. For this reason, we implement the radial tour as the default for the manual tour. Below we describe the manual tour illustrated in detail. After that, the algorithms for oblique cursor movements in 1- and 2D are covered.

### 3.1.1 Notation

The notation used to describe the algorithm for a 2D radial manual tour is as follows:

- $\mathbf{X}$ , the data, an  $n \times p$  numeric matrix to be projected.
- $\mathbf{A}$ , any orthonormal projection basis,  $p \times d$  matrix, describing the projection from  $\mathbb{R}^p \Rightarrow \mathbb{R}^d$ .
- $k$ , is the index of the manipulation variable or manip var for short.
- $\mathbf{e}$ , a 1D basis vector of length  $p$ , with 1 in the  $k$ -th position and 0 elsewhere.

- $\mathbf{R}$ , the  $d + 1$ -D rotation matrix, performs unconstrained 3D rotations within the manip space,  $\mathbf{M}$ .
- $\theta$ , the angle of in-projection rotation, for example, on the reference axes;  $c_\theta, s_\theta$  are its cosine and sine.
- $\phi$ , the angle of out-of-projection rotation, into the manip space;  $c_\phi, s_\phi$  are its cosine and sine. The initial value for animation purposes is  $\phi_1$ .
- $\mathbf{U}$ , the axis of rotation for out-of-projection rotation orthogonal to  $\mathbf{e}$ .
- $\mathbf{Y} = \mathbf{X} \times \mathbf{A}$ , the resulting data projection through the manip space,  $\mathbf{M}$ , and rotation matrix,  $\mathbf{R}$ .

The algorithm operates entirely on projection bases and incorporates the data only when making the projected data plots in light of efficiency.

### 3.1.2 Steps

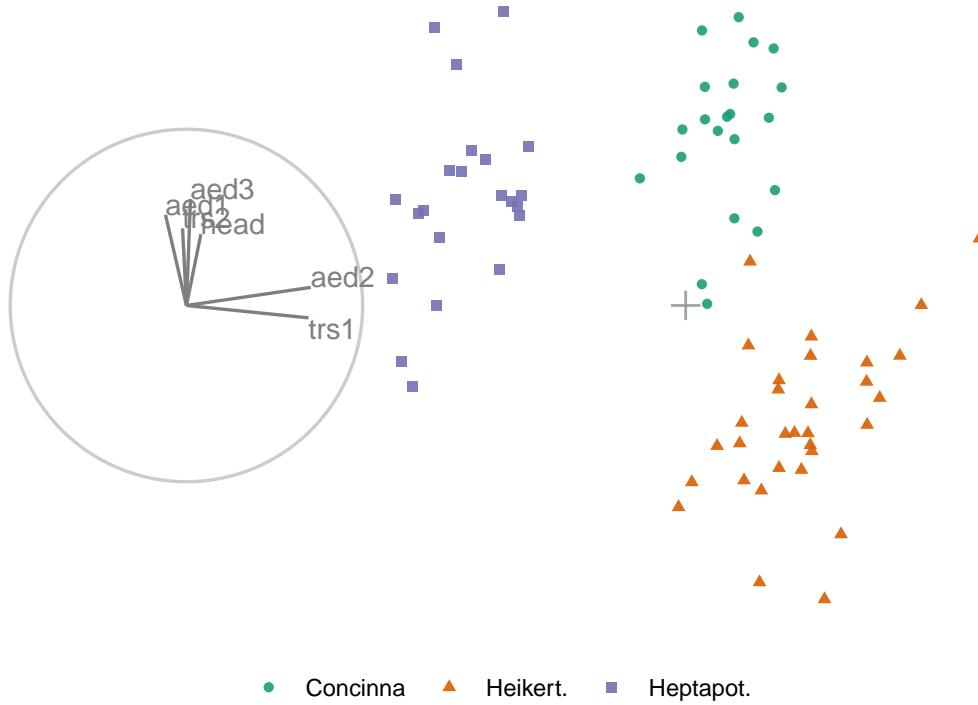
#### Step 0) Setup

The flea data ([Lubischew \(1962\)](#)), available in the `tourr` package ([Wickham et al. 2011](#)), is used to illustrate the algorithm. The data contains 74 observations of six variables, physical measurements of flea beetles. Each observation belongs to one of three species. Each variable is normalized to a common range of  $[0, 1]$ .

An initial 2D projection basis must be provided. A suggested way to start is to identify an interesting projection using a projection pursuit guided tour. The holes index is used to find a 2D projection of the flea data, which shows three separated species groups. Figure 3.1 shows the initial projection of the data. On the left, a biplot ([Gabriel 1971](#)) illustrates the projection basis ( $\mathbf{A}$ ), showing each variable's magnitude and direction of contribution to the projection. The right side shows the projected data,  $\mathbf{Y}_{[n, 2]} = \mathbf{X}_{[n, p]} \mathbf{A}_{[p, 2]}$ . The color and shape of points are mapped to the flea species.

#### Step 1) Choose manip variable

In figure 3.1 the contribution of the variables `tars1` and `aede2` mostly contrast the contribution of the other four variables. These two variables combined contribute in the



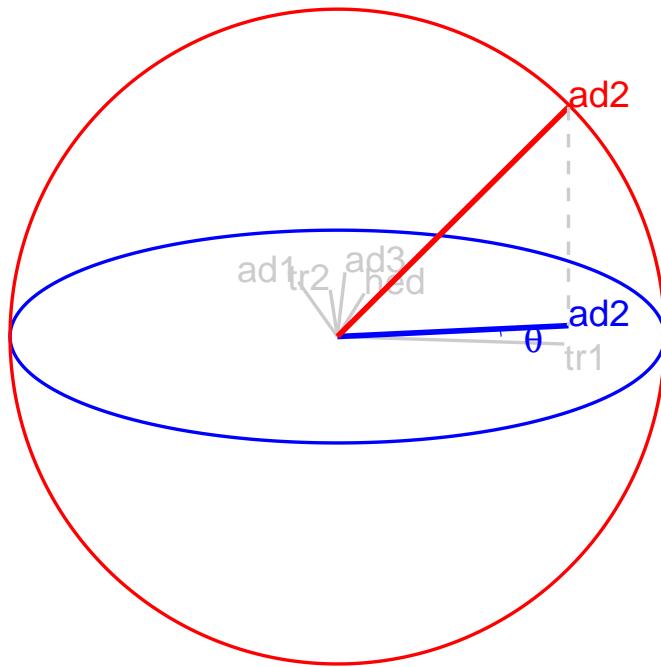
**Figure 3.1:** Biplot of the initial 2D projection of normalized flea data. The basis is depicted on the left and, the resulting data projection is on the right. The color and shape of data points are mapped to the flea beetle species. The basis was produced by a projection pursuit guided tour with the holes index. The contribution of the variables ‘aede2’ and ‘tars1’ approximately contrasts the other variables. The visible structure in the projection is the three species clusters.

direction where the purple cluster is separated from the other two clusters. The variable `aede2` is selected as the manip var, the variable whose contribution is being changed. The aspect being explored is: how important is this variable to separating the clusters in this projection?

### Step 2) Create the 3D manip space

Initialize a zero vector,  $\mathbf{e}$ , of length  $p$ , and set the  $k$ -th element to 1. In the example data, `aede2` is the fifth variable in the data, so  $k = 5$ , set  $e_5 = 1$ . Append this vector to the current basis as the third column. Use a Gram-Schmidt process to orthonormalize the coordinate basis vector on the original 2D projection to describe a 3D manip space,  $\mathbf{M}$ .

$$e_k \leftarrow 1$$



**Figure 3.2:** Illustration of a 3D manip space, the projection plane is shown as a blue circle extending into and out of the display. A manipulation direction is initialized, the red circle, orthogonal to the projection plane. This allows the selected variable, ‘aede2’, to change its contribution to the projection plane. The other variables’ contributions rotate into this space, preserving the orthogonal basis, but are omitted in the manipulation dimension for simplicity.

$$\mathbf{e}_{[p, 1]}^* = \mathbf{e} - \langle \mathbf{e}, \mathbf{A}_1 \rangle \mathbf{A}_1 - \langle \mathbf{e}, \mathbf{A}_2 \rangle \mathbf{A}_2$$

$$\mathbf{M}_{[p, 3]} = (\mathbf{A}_1, \mathbf{A}_2, \mathbf{e}^*)$$

The manip space provides a 3D projection from  $p$ -dimensional space, where the coefficient of the manip var can range completely between  $[0, 1]$ . This 3D space serves as the medium to rotate the projection basis relative to the selected manipulation variable. Figure 3.2 illustrates this 3D manip space with the manip var highlighted. This representation is produced by calling the `view_manip_space()` function. This diagram is purely used to help explain the algorithm.

### Step 3) Defining a 3D rotation

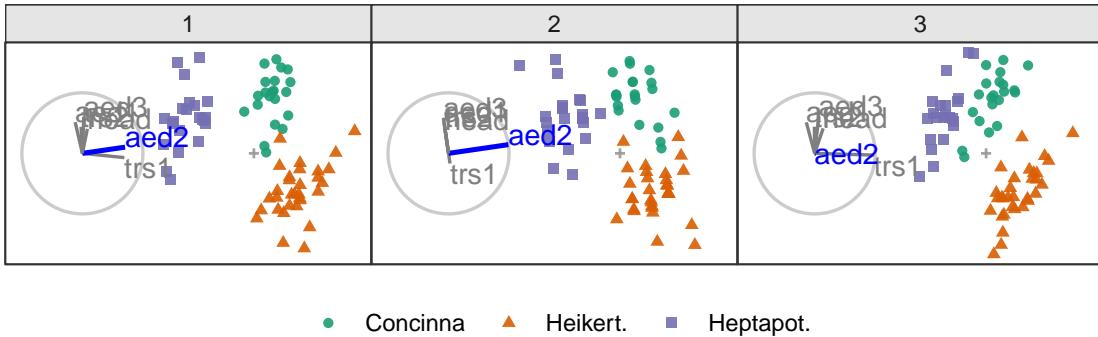
The basis vector corresponding to the manip var (red line in Figure 3.2) can be operated like a lever anchored at the origin. This manual control process rotates the manip variable into and out of the 2D projection (Figure 3.3). As the variable contribution is controlled, the manip space turns, and the projection onto the horizontal plane correspondingly changes. This is a manual tour. Generating a sequence of values for the rotation angles produces a path for the rotation of the manip space.

For a radial tour, fix  $\theta$ , the angle describing rotation within the horizontal projection plane, and compute a sequence for  $\phi$ , defining movement out of this plane. This will change  $\phi$  from the initial value,  $\phi_1$ , the angle between  $\mathbf{e}$  and its shadow in  $\mathbf{A}$ , to a maximum of 0 (manip var entirely in projection), then to a minimum of  $\pi/2$  (manip var orthogonal to the projection), before returning to  $\phi_1$ .

Rotations in 3D can be defined by the axes they pivot on. Rotation within the projection,  $\theta$ , is rotation around the  $Z$ -axis. Out-of-projection rotation,  $\phi$ , is the rotation around an axis on the  $XY$  plane,  $\mathbf{U}$ , orthogonal to  $\mathbf{e}$ . Given these axes, the rotation matrix,  $\mathbf{R}$ , can be written as follows, using Rodrigues' rotation formula (originally published in [Rodrigues \(1840\)](#)):

$$\begin{aligned} \mathbf{R}_{[3, 3]} &= \mathbf{I}_3 + s_\phi \mathbf{U} + (1 - c_\phi) \mathbf{U}^2 \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & c_\theta s_\phi \\ 0 & 0 & s_\theta s_\phi \\ -c_\theta s_\phi & -s_\theta s_\phi & 0 \end{bmatrix} + \begin{bmatrix} -c_\theta(1 - c_\phi) & s_\theta^2(1 - c_\phi) & 0 \\ -c_\theta s_\theta(1 - c_\phi) & -s_\theta^2(1 - c_\phi) & 0 \\ 0 & 0 & c_\phi - 1 \end{bmatrix} \\ &= \begin{bmatrix} c_\theta^2 c_\phi + s_\theta^2 & -c_\theta s_\theta(1 - c_\phi) & -c_\theta s_\phi \\ -c_\theta s_\theta(1 - c_\phi) & s_\theta^2 c_\phi + c_\theta^2 & -s_\theta s_\phi \\ c_\theta s_\phi & s_\theta s_\phi & c_\phi \end{bmatrix} \end{aligned}$$

where



**Figure 3.3:** Select frames of a radial tour manipulating ‘aede2’: (1) original projection, (2) full contribution, (3) zero contribution. After zeroing the contribution the animation continues to return to the initial contribution.

$$\begin{aligned} \mathbf{U} &= (u_x, u_y, u_z) = (s_\theta, -c_\theta, 0) \\ &= \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -c_\theta \\ 0 & 0 & -s_\theta \\ c_\theta & s_\theta & 0 \end{bmatrix} \end{aligned}$$

#### Step 4) Creating an animation of the radial rotation

The steps outlined above can be used to create any arbitrary rotation in the manip space. To use this for sensitivity analysis, the radial rotation is built into an animation where the manip var is rotated fully into the projection, completely out, and then back to the initial value. This involves varying  $\phi$  between 0 and  $\pi/2$ , call the steps  $\phi_i$ .

1. Set the initial value of  $\phi_1$  and  $\theta$ :  $\phi_1 = \cos^{-1} \sqrt{A_{k1}^2 + A_{k2}^2}$ ,  $\theta = \tan^{-1} \frac{A_{k2}}{A_{k1}}$ . Where  $\phi_1$  is the angle between  $\mathbf{e}$  and its shadow in  $\mathbf{A}$ .
2. Set an angle increment ( $\Delta\phi$ ) that sets the step size for the animation, to rotate the manip var into and out of the projection. The package **spinifex** uses an angle increment, rather than number of frames to control the movement to be consistent with the tour algorithm as implemented in **tourr**.
3. Step towards 0, where the manip var is entirely in the projection plane.

4. Step towards  $\pi/2$ , where the manip variable has no contribution to the projection.
5. Step back to  $\phi_1$ .

In each of the steps 3-5, a small step may be added to ensure that the endpoints of  $\phi(0, \pi/2, \phi_1)$  is reached.

### **Step 5) Projecting the data**

The operation of a manual tour is defined on the projection bases. Only when the data plot needs to be made, the data is projected through the relevant basis.

$$\mathbf{Y}_{[n, 3]}^{(i)} = \mathbf{X}_{[n, p]} \mathbf{M}_{[p, 3]} \mathbf{R}_{[3,3]}^{(i)}$$

where  $\mathbf{R}_{[3,3]}^{(i)}$  is the incremental rotation matrix, using  $\phi_i$ . To make the data plot, use the first two columns of  $\mathbf{Y}$ . Show the projected data for each frame in sequence to form an animation.

Tours are typically viewed as an animation. The animation of this tour can be viewed online on [GitHub](#).

## **3.2 Oblique cursor movement**

In a move abbreviated way, we can think about the algorithm for 1D and 2D oblique manual tours as:

---

**Algorithm 1** 1D oblique manual tour from cursor movement

---

**Require:**

- $\mathbf{A}$  is a 1D basis defining the current projection;
- $\mathbf{e} = e_k$  is a  $p$ -dimensional vector of zeros with a 1 in the  $k$ -th position, the manipulation variable;
- $\text{dist}_x$  = horizontal distance of the cursor movement;
- $\text{dist}_y$  = vertical distance of the cursor movement;

**Ensure:**  $\|\mathbf{A} - \mathbf{e}\| > \text{tolerance}$

Initialize the manipulation space, Gram-Schmidt process orthonormalizing  $\mathbf{e}$  on  $\mathbf{A}$ :

$$\mathbf{M} \leftarrow \mathbf{e} - \langle \mathbf{e}, \mathbf{A} \rangle \mathbf{A}$$

Initialize change in magnitude,  $\phi$ , as a function of the horizontal cursor movement:

$$\phi \leftarrow \frac{\text{dist}_x}{\text{size of plot}_x}, \quad c_\phi = \cos(\phi), \quad s_\phi = \sin(\phi)$$

**for**  $\phi_i$  in  $[0, \phi]$  by an interpolation increment step size **do**

    Rotate the manipulation space:

$$\mathbf{M} \leftarrow \mathbf{M} \cdot \begin{pmatrix} c_{\phi_i} & s_{\phi_i} \\ -s_{\phi_i} & c_{\phi_i} \end{pmatrix}$$

    Append an array,  $\mathbf{Y}$ , premultiplying the data by each interpolated basis of the manipulation space:

$$\mathbf{Y}_{npi} \leftarrow \mathbf{X}_{np} \cdot \mathbf{M}_{p1:d}$$

    (Alternatively, store the array of the interpolated bases for a more compact format and premultiply data when needed.)

**end for**

---

---

**Algorithm 2** 2D oblique manual tour from cursor movement

---

**Require:**

- $\mathbf{A}$  is a 1D basis defining the current projection;
- $\mathbf{e} = e_k$  is a  $p$ -dimensional vector of zeros with a 1 in the  $k$ -th position, the manipulation variable;
- $\text{dist}_x$  = horizontal distance of the cursor movement;
- $\text{dist}_y$  = vertical distance of the cursor movement;

**Ensure:**  $\|\mathbf{A} - \mathbf{e}\| > \text{tolerance}$

Initialize the manipulation space, Gram-Schmidt process orthonormalizing  $\mathbf{e}$  on  $\mathbf{A}$ :

$$\mathbf{M} \leftarrow \mathbf{e} - \langle \mathbf{e}, \mathbf{A}_1 \rangle \mathbf{A}_1 - \langle \mathbf{e}, \mathbf{A}_2 \rangle \mathbf{A}_2$$

Initialize change in magnitude,  $\phi$ :

$$\phi \leftarrow \frac{\sqrt{\text{dist}_x^2 + \text{dist}_y^2}}{\text{size of plot}_x}, \quad c_\phi = \cos(\phi), \quad s_\phi = \sin(\phi)$$

Initialize change in angle,  $\theta$ :

$$c_\theta = \cos(\theta) \leftarrow \frac{\text{dist}_x}{\sqrt{\text{dist}_x^2 + \text{dist}_y^2}}$$

$$s_\theta = \sin(\theta) \leftarrow \frac{\text{dist}_y}{\sqrt{\text{dist}_x^2 + \text{dist}_y^2}}$$

**for**  $\phi_i$  in  $[0, \phi]$  by an interpolation increment step size **do**

    Rotate the manipulation space:

$$\mathbf{M} \leftarrow \mathbf{M} \cdot \begin{pmatrix} c_\theta^2 c_\phi + s_\theta^2 & -c_\theta s_\theta (1 - c_\phi) & -c_\theta s_\phi \\ -c_\theta s_\theta (1 - c_\phi) & s_\theta^2 c_\phi + c_\theta^2 & -s_\theta s_\phi \\ c_\theta s_\phi & s_\theta s_\phi & c_\phi \end{pmatrix}$$

    Append an array,  $\mathbf{Y}$ , premultiplying the data by each interpolated basis of the manipulation space:

$$\mathbf{Y}_{npi} \leftarrow \mathbf{X}_{np} \cdot \mathbf{M}_{p1:d}$$

    (Alternatively, store the array of the interpolated bases for a more compact format and premultiply data when needed.)

**end for**

---

### 3.3 Package structure

In addition to facilitating the manual tour, the other primary function is to facilitate the layered composition of tours interoperability with tours from `tourr`. This package attempts to abstract the complexity of dealing with a varying number of frames and replicating the length of arguments. We use a layered composition approach to tours similar to `ggplot2` (Wickham 2016). The resulting displays then be animated with `plotly` (Sievert 2020) or `ganimate` (Pedersen and Robinson 2020). This section describes the functions available in the package, their usage, and how to install and get up and running.

#### 3.3.1 Usage

The penguins data (Gorman, Williams, and Fraser 2014; Horst, Hill, and Gorman 2020), also available in `spinifex`, is used to illustrate the creation of a radial tour, layered composition, and animation. Similar to the flea data, this is exploring the variable contribution that leads to the separation of clusters.

```
library(spinifex)

## Process penguins data
dat <- scale_sd(penguins_na.rm[1:4])
clas <- penguins_na.rm$species

## Start basis and manual tour path
bas <- basis_olda(data = dat, class = clas)
mt_path <- manual_tour(basis = bas, manip_var = 1, data = dat)

## Compose the tour display
ggt <- ggplot(mt_path, angle = .15) +
  proto_point(aes_args      = list(color = clas, shape = clas),
              identity_args = list(alpha = .8, size = 1.5)) +
  proto_basis() +
```

```
proto_origin()

## Animate

animate_plotly(ggt, fps = 5)
```

Composition and animation options are interoperable with **tourr** produced tours. The code below similarly composes a 1D tour from a saved grand tour path.

```
## A grand tour from tourr

library(tourr)

gt_path <- save_history(data = dat, tour_path = grand_tour(), max_bases = 10)

## Compose 1D tour display

ggt2 <- ggtour(gt_path, angle = .15) +
  proto_density(aes_args = list(color = clas, fill = clas)) +
  proto_basis1d() +
  proto_origin1d()

## Animate

animate_plotly(ggt2, fps = 5)
```

### 3.3.2 Functions

Table 3.1 lists the primary functions and their purpose. These are grouped into four classes: processing the data, production of tour path, the composition of the tour display, and its animation.

### 3.3.3 Installation

The **spinifex** is available from CRAN. The following code will help to get up and running:

**Table 3.1:** Summary of primary functions.

| Family     | Function           | Related to          | Description   |
|------------|--------------------|---------------------|---|
| processing | scale_01/sd        | -                   | scale each column to [0,1]/std dev away from the mean   |
| processing | basis_pca/olda/... | Rdimtools::do.*     | basis of orthogonal component spaces                    |
| processing | basis_half_circle  | -                   | basis with uniform contribution across half of a circle |
| processing | basis_guided       | tourr::guided_tour  | silently return the basis from a guided tour            |
| tour path  | manual_tour        | -                   | basis and interpolation information for a manual tour   |
| tour path  | save_history       | tourr::save_history | silent, extended wrapper returning other tour arrays    |
| display    | ggtour             | ggplot2::ggplot     | canvas and initialization for a tour animation          |
| display    | proto_point/text   | geom_point/text     | adds observation points/text                            |
| display    | proto_density/2d   | geom_density/2d     | adds density curve/2d contours                          |
| display    | proto_hex          | geom_hex            | adds hexagonal heatmap of observations                  |
| display    | proto_basis/1d     | -                   | adds adding basis visual in a unit-circle/-rectangle    |
| display    | proto_origin/1d    | -                   | adds reference mark in the center of the data           |
| display    | proto_default/1d   | -                   | wrapper for proto_* point + basis + origin              |
| display    | facet_wrap_tour    | ggplot2::facet_wrap | facets on the levels of variable                        |
| display    | append_fixed_y     | -                   | add/overwrite a fixed vertical position                 |
| animation  | animate_plotly     | plotly::ggplotly    | render as an interactive html widget                    |
| animation  | animate_gganimate  | gganimate::animate  | render as a gif, mp4, or other video formats            |
| animation  | filmstrip          | -                   | static ggplot faceting on the frames of the animation   |

```
# Installation:

install.package("spinifex") ## Install from CRAN

library("spinifex") ## Load into session


# Getting started:

## Shiny app for visualizing basic application

run_app("intro")

## View the code vignette

vignette("getting_started_with_spinifex")

## More about proto_* functions

vignette("ggproto_api")
```

## 3.4 Use cases

Wang et al. (2018) introduce a new tool, PDFSense, to visualize the sensitivity of hadronic experiments to nucleon structure. The parameter-space of these experiments lies in 56 dimensions, and are approximated as the ten first principal components.

Cook, Laa, and Valencia (2018) illustrate using the grand tour to explore this component reduced space. Tours can better resolve the shape of clusters, intra-cluster detail, and lead to better outlier detection than PDFSense & TFEP (TensorFlow embedded projections)

or traditional static embeddings. Building on this, the manual tour is used to examine the sensitivity of structure in a projection to different parameters. Bases from the grand tour identified in the previous work are used as the initial basis.

The data has a hierarchical structure with top-level clusters; DIS, VBP, and jet. Each cluster is a particular class of experiments, each with many experimental datasets. In consideration of data occlusion, we conduct manual tours on subsets of the DIS and jet clusters. This explores the structure's sensitivity to each of the variables in turn, and we present the subjectively best and worst variable to manipulate for identifying dimensionality of the clusters and describing the range of the clusters.

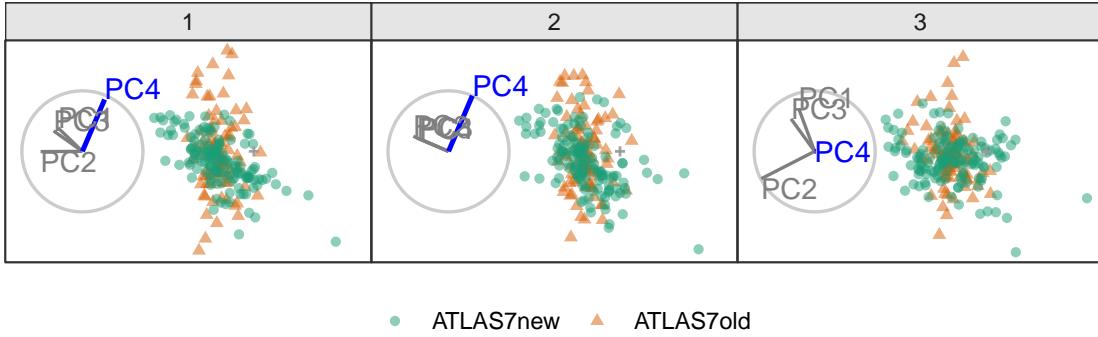
### 3.4.1 Jet cluster

The jet cluster resides in a smaller dimensionality than the full set of experiments, with four principal components explaining 95% of the variation in the cluster ([Cook, Laa, and Valencia 2018](#)). The data within this 4D embedding is a further subset to ATLAS7old and ATLAS7new, focuses on two groups that occupy different parts of the subspace. Radial manual tours controlling contributions from PC4 and PC3 are shown in Figures [3.4](#) and [3.5](#), respectively. The difference in shape can be interpreted as the experiments probing different phase spaces. Back-transforming the principal components to the original variables can be done for a more detailed interpretation.

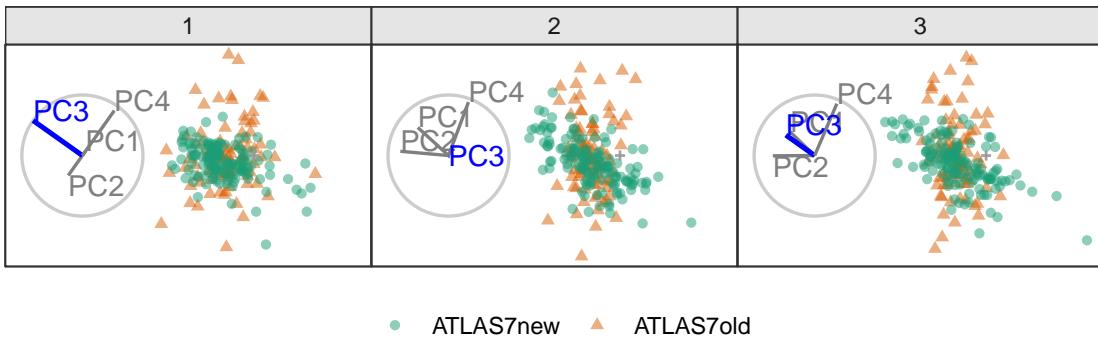
When PC4 is removed from the projection (Figure [3.4](#)), the difference between the two groups is removed, indicating that PC4 is essential for separating types of experiments. However, eliminating PC3 from the projection (Figure [3.5](#)) does not affect the structure, meaning PC3 is not essential for distinguishing experiments. Animations for the remaining PCs can be viewed at the following links: [PC1](#), [PC2](#), [PC3](#), and [PC4](#). It can be seen that only PC4 is vital for viewing the difference in these two experiments.

### 3.4.2 DIS cluster

Following [Cook, Laa, and Valencia \(2018\)](#), PCA is used to explore the DIS cluster, and the first six principal components are used, which contain 48% of the full sample variation. The contributions of PC6 and PC2 are explored in Figures [3.6](#) and [3.7](#), respectively. Three



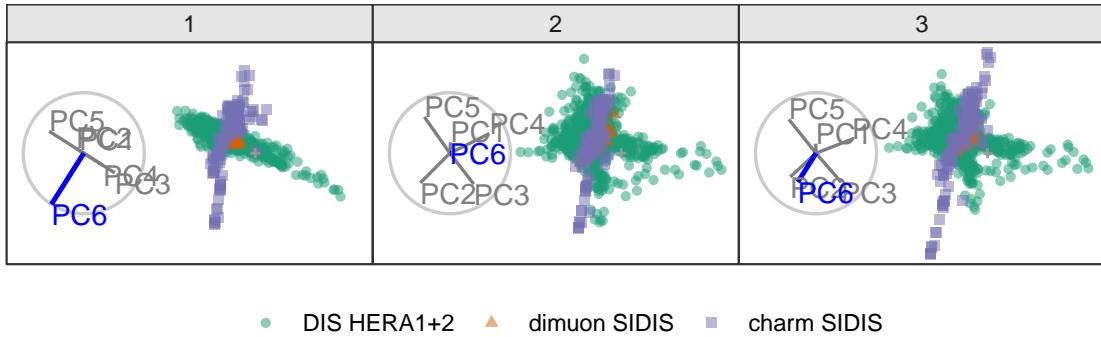
**Figure 3.4:** Select frames from a radial tour of PC4 within the jet cluster, with color indicating experiment type: ATLAS7new (green) and ATLAS7old (orange). When PC4 is removed from the projection (frame 10), there is little difference between the clusters, suggesting that PC4 is vital for distinguishing the experiments.



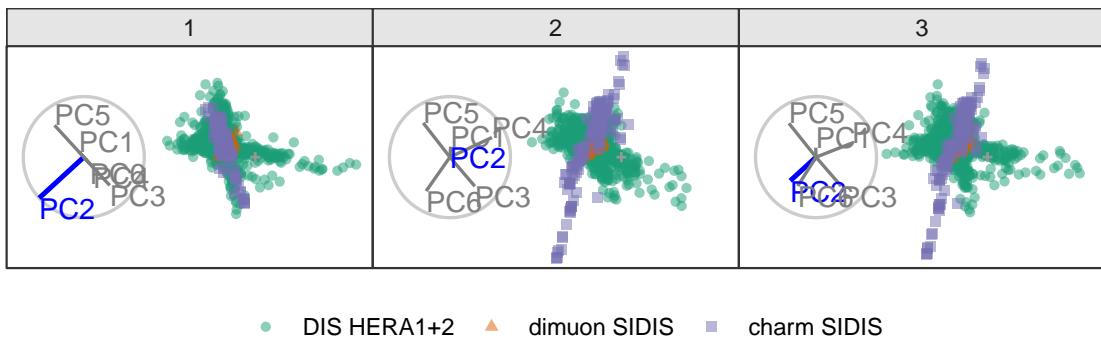
**Figure 3.5:** Frames from the radial tour manipulating PC3 within the jet cluster, with color indicating experiment type: ATLAS7new (green) and ATLAS7old (orange). When the contribution from PC3 is changed, there is little change in the separation of the clusters, suggesting that PC3 is not important for distinguishing the experiments.

experiments are examined: DIS HERA1+2 (green), dimuon SIDIS (purple), and charm SIDIS (orange).

Both PC2 and PC6 contribute to the projection similarly. When PC6 is rotated into the projection, variation in the DIS HERA1+2 is greatly reduced. When PC2 is removed from the projection, dimuon SIDIS becomes more distinct. Even though both variables contribute similarly to the original projection, their contributions have quite different effects on the structure of each cluster and the distinction between clusters. Animations of all of the principal components can be viewed from the links: [PC1](#), [PC2](#), [PC3](#), [PC4](#), [PC5](#), and [PC6](#).



**Figure 3.6:** Select frames from a radial tour exploring the sensitivity that  $PC_6$  has on the structure of the DIS cluster, with color indicating experiment type: DIS HERA1+2 (green), dimuon SIDIS (purple), and charm SIDIS (orange). DIS HERA1+2 is distributed in a cross-shaped plane, and charm SIDIS occupies the center of this cross, and dimuon SIDIS is a linear cluster crossing DIS HERA1+2. As the contribution of  $PC_6$  is increased, DIS HERA1+2 becomes almost singular in one direction (frame 5), indicating that this cluster has minimal variability in the direction of  $PC_6$ .



**Figure 3.7:** Frames from the radial tour exploring the sensitivity  $PC_2$  to the structure of the DIS cluster, with color indicating experiment type: DIS HERA1+2 (green), dimuon SIDIS (purple), and charm SIDIS (orange). As the contribution of  $PC_2$  is decreased, dimuon SIDIS becomes more distinguishable from the other two clusters, indicating that the absence of  $PC_2$  is vital for separating this cluster from the others.

### 3.5 Discussion

Dynamic linear projections of numeric multivariate data, tours, play an important role in data visualization; they extend the dimensionality of visuals to peek into high-dimensional data and parameter spaces. This research has taken the manual tour algorithm, specifically the radial rotation, used in GGobi ([Swayne et al. 2003](#)) to interactively rotate a variable into or out of a 2D projection and modified it to create an animation that performs the same task. It is most helpful in examining the importance of variables and how the structure in the projection is sensitive or not to specific variables. This functionality is made available in the package **spinifex**. Which also extends the geometric display and export formats interoperable with the **tourr** package.

The original conception of the grand tour was motivated by problems in physics ([Asimov 1985](#)) Thus, the use case was fitting. The radial tour was used to explore the sensitivity of variable contributions to the separation of different experiments in high-energy particle physics. These tools can be applied quite broadly to many multivariate data analysis problems.

The manual tour is constrained because the effect of one variable depends on the contributions of other variables in the manip space. However, this can simplify a projection by removing variables without affecting the visible structure. Defining a manual rotation in high dimensions is possible using Givens rotations and Householder reflections as outlined in [Buja et al. \(2005\)](#). This would provide more flexible manual rotation but more difficult for a user because they have the choice (too much choice) of which directions to move.



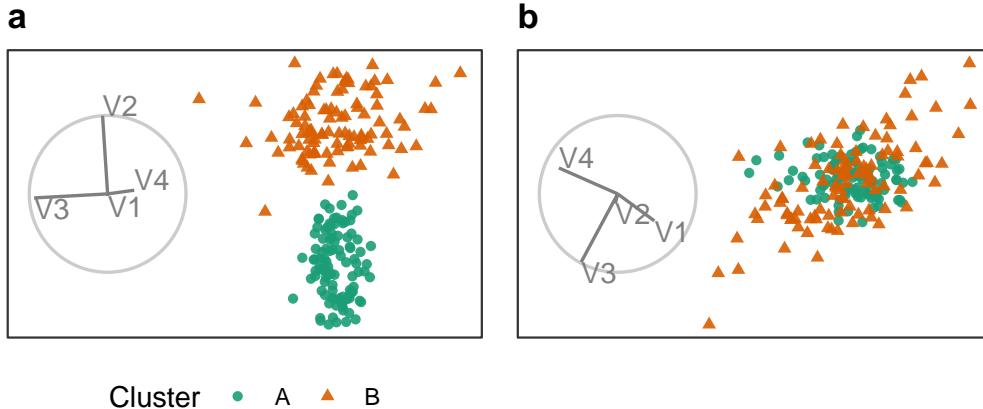
## Chapter 4

# A Study on the Benefit of a User-Controlled Radial Tour for Variable Attribution of Structure

The previous chapter introduced the package **spinifex**, which provides the means to perform radial tours. However, there is no evidence suggesting that the radial tour's user-controlled steering leads to a better perception than traditional methods. Therefore, this chapter discusses the user study to elucidate the efficacy of the radial tour.

Chapters 1 and 2 introduced PCA, the grand tour, and the radial tour. This chapter describes a within-participants user study evaluating the efficacy of these techniques. A supervised classification task is designed where participants evaluate variable attribution of the separation between two classes. An accuracy measure is defined as a response variable. Data were collected from 108 crowdsourced participants, who performed two trials with each visual for 648 trials in total.

The user influence over a basis is crucial to testing variable sensitivity to the structure visible in projection and uniquely available in the radial tour. If the contribution of a variable is reduced, and the feature disappears, then it is said that the variable is sensitive to that structure. For example, Figure 4.1 shows two projections of simulated data. Panel



**Figure 4.1:** Illustration of cluster separation. Panel (a) shows clear separation in the direction of  $V_2$  and no separation in the direction of  $V_3$ . While the contributions of  $V_1$  and  $V_4$  are relatively small and contributed little to this orientation. Panel (b) has a minimal contribution from  $V_2$ , and no separation between the cluster means is resolved.

(a) has identified separation between the two clusters and primarily in the direction of the contribution from  $V_2$ . Many other bases do not reveal this cluster separation, such as panel (b) with a small contribution from  $V_2$ . Because of this, it is said that  $V_2$  is sensitive to the separation of the clusters.

Understanding which variables to use is also important to statistical modeling and their interpretations. Models are becoming increasingly complex, and the nonlinear interactions of their terms cause an opaqueness to model interpretability. Exploratory Artificial Intelligence (XAI, [Adadi and Berrada 2018](#); [Arrieta et al. 2020](#)) is an emerging field that extends the interpretability of such models. Multivariate data visualization is essential for exploring features spaces and communicating interpretations of models ([Biecek 2018](#); [Biecek and Burzykowski 2021](#); [Wickham, Cook, and Hofmann 2015](#)).

This chapter is structured as follows. Section 4.1 describes the implementation of the user study; its visual methods, experimental factors, task, accuracy measure used standardization, and randomization of the experiment. The results of the study are discussed in Section 4.2. Conclusions and potential future directions are discussed in Section 4.3. An accompanying application and extended analysis are provided in the appendix under Section ??.

## 4.1 User study

An experiment was constructed to assess the performance of the radial tour relative to the grand tour and PCA for interpreting the variable attribution contributing to separation between two clusters. Data were simulated across three experimental factors: location of the cluster separation, cluster shape, and data dimensionality. Participant responses were collected using a web application and crowdsourced through prolific.co, ([Palan and Schitter 2018](#)) an alternative to MTurk.

### 4.1.1 Objective

PCA will be used as a baseline for comparison as it is the most commonly used linear embedding. It will use static, discrete jumps between orthogonal components. The grand tour will act as a secondary control that will help evaluate the benefit of observation trackability between nearby animation frames but without user-control of its path. Lastly, the radial tour will be compared, which benefits from the continuity of animation and user control of the basis.

Then for some subset of tasks, we expect to find that the radial tour performs most accurately. Conversely, we are less sure about the accuracy of such limited grand tours as there is no objective function in selecting the bases; it is possible that the random selection of the target bases altogether avoids bases showing cluster separation. However, given that the data dimensionality is modest, it seems plausible that the grand tour coincidentally regularly crossed bases with the correct information for the task.

Experimental factors and the definition of an accuracy measure are given below. The null hypothesis can be stated as:

$$H_0 : \text{accuracy does not change across the visual methods}$$

$$H_\alpha : \text{accuracy does change across the visual methods}$$

### 4.1.2 Experimental factors

In addition to the visual method, data are simulated across three experimental factors. First, the *location* of the separation between clusters is controlled by mixing a signal and a noise variable at different ratios. Secondly, the *shape* of the clusters reflect varying distributions of the data. And third, the *dimensionality* of the data is also tested. The levels within each factor are described below, and Figure 4.2 gives a visual representation.

The *location* of the separation between the clusters is at the heart of the measure. It would be good to test a few varying levels. To test the sensitivity, a noise, and signal-containing variable are mixed. The separation between clusters are mixed at the following percentages: 0/100% (not mixed), 33/66%, 50/50% (evenly mixed).

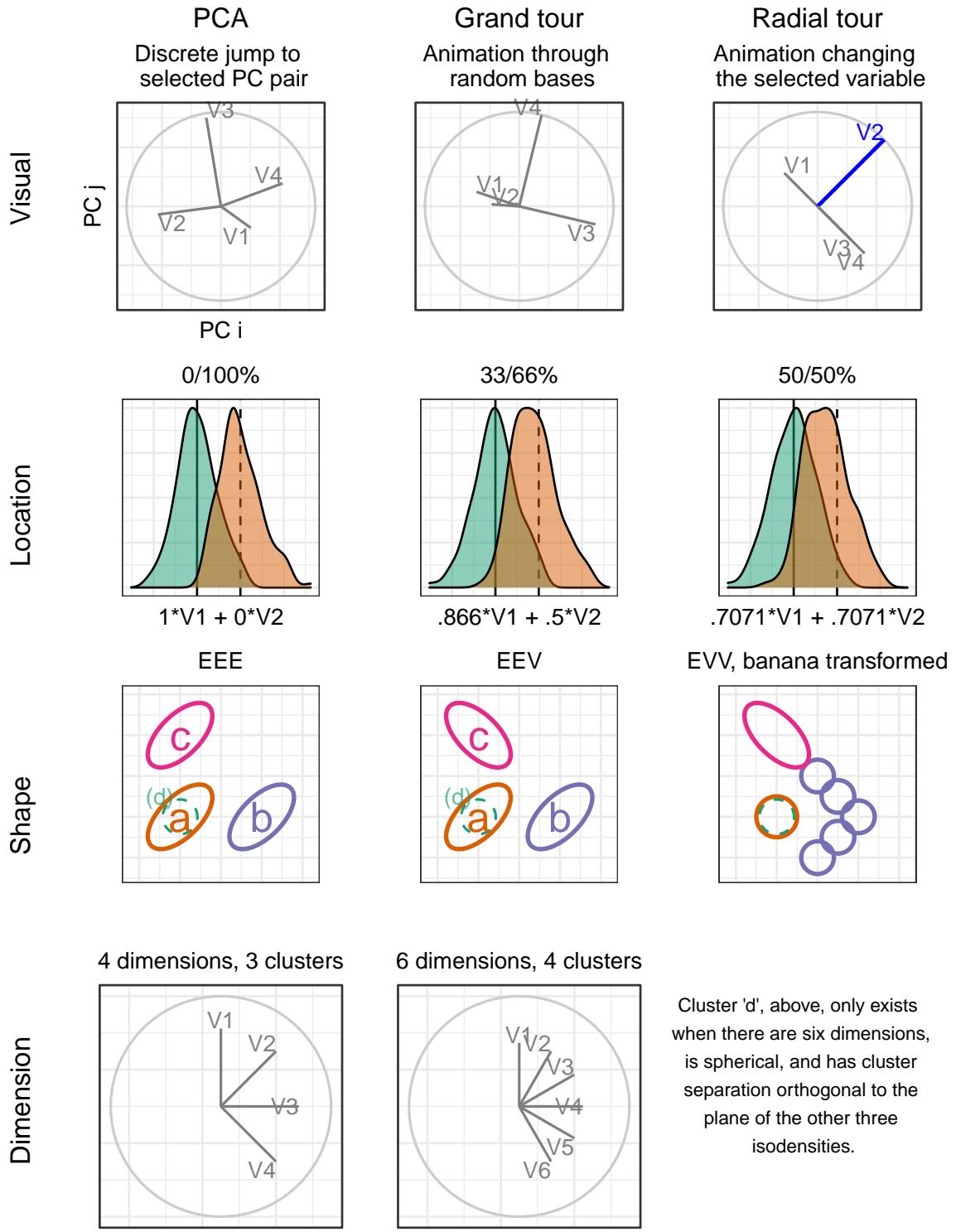
In selecting the *shape* of the clusters, the convention given by Scrucca et al. (2016) is followed. They describe 14 variants of model families containing three clusters. The model family name is the abbreviation of the cluster's respective volume, shape, and orientation. The levels are either *Equal* or *Vary*. The models EEE, EEV, and EVV are used. For instance, in the EEV model, the volume and shape of clusters are constant, while the shape's orientation varies. The EVV model is modified by moving four-fifths of the data out in a “>” or banana-like shape.

*Dimensionality* is tested at two modest levels: four dimensions containing three clusters and six with four clusters. Such modest dimensionality is required to limit the difficulty and search space to make the task realistic for crowdsourcing.

### 4.1.3 Task and evaluation

With our hypothesis formulated and data at hand, let us turn our attention to the task and how to evaluate it. Regardless of the visual method, the display elements are held constant, shown as a 2D scatterplot with a biplot (Gabriel 1971) to its left. A biplot is a visual depiction of the variable contributions from the basis inscribed in a unit circle. Observations were supervised with cluster membership mapped to (colorblind safe) colors and shape.

## Levels of the experimental factors



**Figure 4.2:** Illustration of the experimental factors, the parameter space of the independent variables.

Participants were asked to “check any/all variables that contribute more than average to the cluster separation green circles and orange triangles,” which was further explained in the explanatory video as “mark any and all variable that carries more than their fair share of the weight, or one quarter in the case of four variables.”

The instructions were iterated several times in the video was: 1) use the input controls to find a frame that contains separation between the clusters of green circles and orange triangles, 2) look at the orientation of the variable contributions in the gray circle (biplot axes orientation), and 3) select all variables that contribute more than uniformed distributed cluster separation in the scatterplot. Independent with experimental level, participants were limited to 60 seconds for each evaluation of this task. This restriction did not impact many participants as the 25th, 50th, 75th quantiles of the response time were about 7, 21, and 30 seconds, respectively.

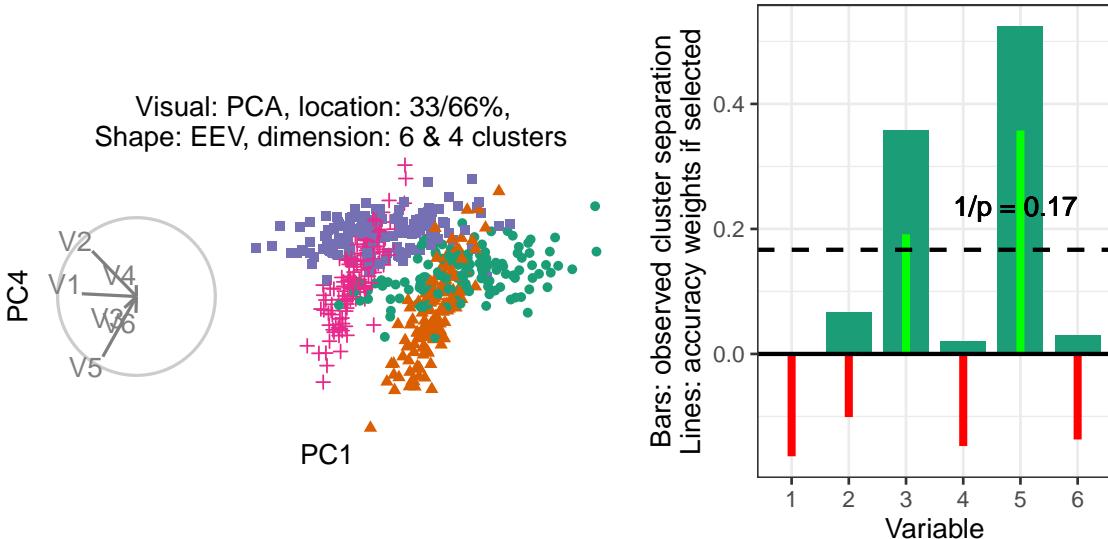
The accuracy measure of this task was designed with a couple of features in mind. 1) symmetric about the expected value, that is, without preference to under- or over-guessing. 2) heavier than linear weight with an increasing difference from the expected value. The following measure is defined for evaluating the task.

Let the data  $\mathbf{X}_{n, p, k}$  be a simulation containing clusters of observations of different distributions. Where  $n$  is the number of observations,  $p$  is the number of variables, and  $k$  indicates the observation’s cluster. Cluster membership is exclusive; an observation cannot belong to more than one cluster.

The weights,  $w$ , is a vector, the variable-wise difference between the mean of two clusters of less  $1/p$ , the expected cluster separation if it were uniformly distributed. Accuracy,  $A$  is defined as the signed square of these weights if selected by the participant. Participant responses are a logical value for each variable — whether or not the participant thinks each variable separates the two clusters more than uniformly distributed separation.

$$w_j = \frac{(\bar{X}_{\cdot, j=1, k=1} - \bar{X}_{\cdot, 1, 2}, \dots, (\bar{X}_{\cdot, p, 1} - \bar{X}_{\cdot, p, 2})}{\sum_{j=1}^p (|\bar{X}_{\cdot, j, k=1} - \bar{X}_{\cdot, j, 2}|)} - \frac{1}{p}$$

$$A = \sum_{j=1}^p I(j) \cdot sign(w_j) \cdot w^2$$



**Figure 4.3:** Illustration of how accuracy is measured. (L), Scatterplot and biplot of PC1 by PC4 of a simulated data set (R) illustrates cluster separation between the green circles and orange triangles. Bars indicate observed cluster separation, and (red/green) lines show the accuracy weight of the variable if selected. The horizontal dashed line is  $1/p$ , the expected value of cluster separation. The accuracy weights equal the signed square of the difference between each variable value and the dashed line.

Where  $I(j)$  is the indicator function, the binary response for variable  $j$ . Figure 4.3 shows one projection of a simulation with its observed variable separation (wide bars), expected uniform separation (dashed line), and accuracy if selected (thin lines).

#### 4.1.4 Visual design standardization

The visual methods are tested within-participant, with each visual being evaluated twice by each participant. The order that experimental factors are experienced is randomized with the assignment, as illustrated in Figure 4.4. Below discusses the design standardization and unique input associated with each visual.

The visualization methods were standardized wherever possible. Data were displayed as 2D scatterplots with biplots. All aesthetic values (colors, shapes, sizes, absence of legend, and axis titles) were constant. Variable contributions were always shown left of the scatterplot embeddings with their aesthetic values consistent. What did vary between visuals were their inputs.

PCA allowed users to select between the top four principal components for each axes regardless of the data dimensionality (four or six). Upon changing an axis, the visual would change to the new view of orthogonal components without displaying intermediate bases. There was no user input for the grand tour; users were instead shown a 15-second animation of the same randomly selected path (variables containing cluster separation were shuffled after simulation). Participants could view the same clip up to four times within the time limit. Radial tours allowed participants to select the manipulation variable. The starting basis was initialized to a half-clock design, where the variables were evenly distributed in half of the circle. This design was created to be variable agnostic while maximizing the independence of the variables. Selecting a new variable resets the animation where the new variable is manipulated to a complete contribution, zeroed contribution, and then back to its initial contribution. Animation and interpolation parameters were held constant across grand and radial tour (five frames per second with a step size of 0.1 radians between interpolated frames).

#### 4.1.5 Data simulation

Each dimension is distributed initially as  $\mathcal{N}(0, 1)$ , given the covariance set by the shape factor. Clusters were initially separated by a distance of two before location mixing. Signal variables had a correlation of 0.9 when they had equal orientation and -0.9 when their orientations varied. Noise variables were restricted to zero correlation. Each cluster is simulated with 140 observations and is offset in a variable that did not distinguish previous variables.

Clusters of the EVV shape are transformed to the banana-chevron shape (illustrated in figure 4.2, shape row). Then location mixing is applied by post-multiplying a rotation matrix to the signal variable and a noise variable for the clusters in question. All variables are then standardized by standard deviations away from the mean. The columns are then shuffled randomly.

Each of these replications is then iterated with each level of the visual. For PCA, projections were saved (to png) for each of the 12 pairs of the top four principal components. A grand tour basis path is saved for each dimensionality level. The data from each simulation is

then projected through its corresponding bases path and saved to `gif` file. The radial tour starts at either the four or six-variable “half-clock” basis. A radial tour is then produced for each variable and saved as a `gif`.

#### 4.1.6 Randomized assignment

Now, with simulation and their artifacts in hand, this section covers how the experimental factors are assigned and demonstrate how this is experienced from the participant’s perspective.

The study is sectioned into three periods. Each period is linked to a randomized level of visual and location. The order of dimension and shape are of secondary interest and are held constant in increasing order of difficulty; four then six dimensions and EEE, EVV, then EVV-banana, respectively.

Each period starts with an untimed training task at the simplest remaining experimental levels; location = 0/100%, shape = EEE, and four dimensions with three clusters. This serves to introduce and familiarize participants with input and visual differences. After the training, the participant performs two trials with the same visual and location level across the increasing difficulty of dimension and shape. The plot was removed after 60 seconds, though participants rarely reached this limit.

The order of the visual and location levels is randomized with a nested Latin square where all levels of the visuals are exhausted before advancing to the next level of location. This requires  $3!^2 = 36$  participants to evaluate all permutations of the experimental factors once. This randomization controls for potential learning effects the participant may receive. Figure 4.4 illustrates how an arbitrary participant experiences the experimental factors.

Through pilot studies sampled by convenience (information technology and statistics Ph.D. students attending Monash University), it was estimated that three full evaluations are needed to power the study properly, a total of  $N = 3 \times 3!^2 = 108$  participants.

#### Randomization for the 63<sup>rd</sup> participant

- 1) Set the visual order:  
 $1 + (63 - 1) \text{ mod } 6 =$   
 permutation 4;  
 grand, PCA, & radial
- 2) Set location order:  
 $1 + \text{floor}((63 - 1) / 6) \text{ mod } 36 =$   
 permutation 3; 33/67, 50/50,  
 & 0/100 % noise/signal mix
- Fixed factors:  
 3) Variance-covariance shape  
 increments with each period:  
 EEE, EEV, EVV-banana  
 4) Data dimension is fixed within  
 each period: 4 then 6

| visual order permutations | Period 1 |   |   | Period 2 |   |   | Period 3 |   |   | location order permutations | Period 1 |   |   | Period 2 |   |   | Period 3 |   |   |
|---------------------------|----------|---|---|----------|---|---|----------|---|---|-----------------------------|----------|---|---|----------|---|---|----------|---|---|
|                           | 1        | 2 | 3 | 1        | 2 | 3 | 1        | 2 | 3 |                             | 1        | 2 | 3 | 1        | 2 | 3 | 1        | 2 | 3 |
| 1                         | P        | G | R |          |   |   | 1        | 1 | 2 | 3                           |          |   |   |          |   |   |          |   |   |
| 2                         | P        | R | G |          |   |   | 2        | 1 | 3 | 2                           |          |   |   |          |   |   |          |   |   |
| 3                         | G        | R | P |          |   |   | 3        | 2 | 3 | 1                           |          |   |   |          |   |   |          |   |   |
| 4                         | G        | P | R |          |   |   | 4        | 2 | 1 | 3                           |          |   |   |          |   |   |          |   |   |
| 5                         | R        | P | G |          |   |   | 5        | 3 | 1 | 2                           |          |   |   |          |   |   |          |   |   |
| 6                         | R        | P | G |          |   |   | 6        | 3 | 2 | 1                           |          |   |   |          |   |   |          |   |   |

Increment through all permutations of visual, before incrementing 1 permutation of location

| Period | Evaluation order | Visual <sup>1</sup> | Location <sup>2</sup> | Shape <sup>3</sup> | Dimensions <sup>4</sup> |
|--------|------------------|---------------------|-----------------------|--------------------|-------------------------|
| 1      | Train 1          | Grand               | 0/100                 | EEE                | 4 (3cl)                 |
| 1      | 1                | Grand               | 33/67                 | EEE                | 4 (3cl)                 |
| 1      | 2                | Grand               | 33/67                 | EEE                | 6 (4cl)                 |
| 2      | Train 2          | PCA                 | 0/110                 | EEE                | 4 (3cl)                 |
| 2      | 3                | PCA                 | 50/50                 | EEV                | 4 (3cl)                 |
| 2      | 4                | PCA                 | 50/50                 | EVV                | 6 (4cl)                 |
| 3      | Train 3          | Radial              | 0/100                 | EEE                | 4 (3cl)                 |
| 3      | 5                | Radial              | 0/100                 | Ban                | 4 (3cl)                 |
| 3      | 6                | Radial              | 0/100                 | Ban                | 6 (4cl)                 |

**Figure 4.4:** Illustration of how a hypothetical participant 63 is assigned experimental factors. Each of the six visual order permutations is exhausted before iterating to the next permutation of location order.

#### 4.1.7 Participants

$N = 108$  participants were recruited via prolific.co (Palan and Schitter 2018). Participants are restricted based on their claimed education requiring that they have completed at least an undergraduate degree (some 58,700 of the 150,400 users at the time). This restriction is used on the premise that linear projections and biplot displays will not be regularly used for consumption by general audiences. There is also the implicit filter that Prolific participants must be at least 18 years of age and implicit biases of timezone, location, and language. Participants were compensated for their time at £7.50 per hour, whereas the mean duration of the survey was about 16 minutes. Previous knowledge or familiarity was minimal, as validated in the follow-up survey. The appendix Section B.3.2 contains a heatmap distribution of age and education paneled across preferred pronouns of the participants that completed the survey, who are relatively young, well educated, and slightly more likely to identify as males.

#### 4.1.8 Data collection

Data were recorded in `shiny` application and written to a Google Sheet after each third of the study. Especially at the start of the study, participants experienced adverse network conditions due to the volume of participants hitting the application with modest allocated resources. In addition to this, API read/write limitations further hindered data collection. To mitigate this, the number of participants was throttled and over-collect survey trials until three evaluations were collected for all permutation levels.

The processing steps were minimal. The data were formatted and then filtered to the latest three complete studies of each experimental factor, which should have experienced the least adverse network conditions. The bulk of the studies removed were partial data and a few over-sampled permutations. This brings us to the 108 studies described in the paper, from which models and aggregation tables were built. The post-study surveys were similarly decoded to human-readable format and then filtered to include only those 84 associated with the final 108 studies.

The code, response files, their analyses, and the study application are publicly available at [https://github.com/nspyrison/spinifex\\_study](https://github.com/nspyrison/spinifex_study).

## 4.2 Results

To recap, the primary response variable is accuracy, as defined in Section 4.1.3. The parallel analysis of the log response time is provided in the appendix, Section B.3.3. Two primary data sets were collected; the user study evaluations and the post-study survey. The former is the 108 participants with the experimental factors: visual, location of the cluster separation signal, the shape of variance-covariance matrix, and the dimensionality of the data. Experimental factors and randomization were discussed in section 4.1.2. A follow-up survey was completed by 84 of these 108 people. It collected demographic information (preferred pronoun, age, and education), and subjective measures for each visual (preference, familiarity, ease of use, and confidence).

Below a battery of mixed regression models is built to examine the degree of the evidence and the size of the effects from the experimental factors. Then, Likert plots and rank-sum tests to compare the subjective measures between the visuals.

### 4.2.1 Accuracy

To quantify the contribution of the experimental factors to the accuracy, a mixed-effects models were fit. All models have a random effect term on the participant and the simulation. These terms explain the amount of error that can be attributed to the individual participant's effect and variation due to the random sampling data.

In building a set of models to test, a base model with only the visual term is compared with the full linear model term and progressively interacting an additional experimental factor. The models with three and four interacting variables are rank deficient; there is not enough varying information in the data to explain all interacting terms.

**Fixed effects    Full model**

$$\begin{aligned}
 \alpha & \quad \hat{Y} = \mu + \alpha_i + \mathbf{Z} + \mathbf{W} + \epsilon \\
 \alpha + \beta + \gamma + \delta & \quad \hat{Y} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \epsilon \\
 \alpha \times \beta + \gamma + \delta & \quad \hat{Y} = \mu + \alpha_i \times \beta_j + \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \epsilon \\
 \alpha \times \beta \times \gamma + \delta & \quad \hat{Y} = \mu + \alpha_i \times \beta_j \times \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \epsilon \\
 \alpha \times \beta \times \gamma \times \delta & \quad \hat{Y} = \mu + \alpha_i \times \beta_j \times \gamma_k \times \delta_l + \mathbf{Z} + \mathbf{W} + \epsilon
 \end{aligned}$$

where  $\mu$ , the intercept of the model

$\alpha_i$ , fixed term for visual |  $i \in (\text{pca}, \text{grand}, \text{radial})$

$\beta_j$ , fixed term for location |  $j \in (0/100\%, 33/66\%, 50/50\%)$  noise/signal mixing

$\gamma_k$ , fixed term for shape |  $k \in (\text{EEE}, \text{EEV}, \text{EVV banana})$  model shapes

$\delta_l$ , fixed term for dimension |  $l \in (4 \text{ variables \& 3 cluster}, 6 \text{ variables \& 4 clusters})$

$\mathbf{Z} \sim \mathcal{N}(0, \tau)$ , the error of the random effect of participant

$\mathbf{W} \sim \mathcal{N}(0, v)$ , the error of the random effect of simulation

$\epsilon \sim \mathcal{N}(0, \sigma)$ , the remaining error in the model

**Table 4.1:** Model performance of random effect models regressing accuracy. Complex models perform better in terms of  $R^2$  and RMSE, yet AIC and BIC penalize their large number of fixed effects in favor of the much simpler model containing only the visuals. Conditional  $R^2$  includes error explained by the random effects, while marginal does not.

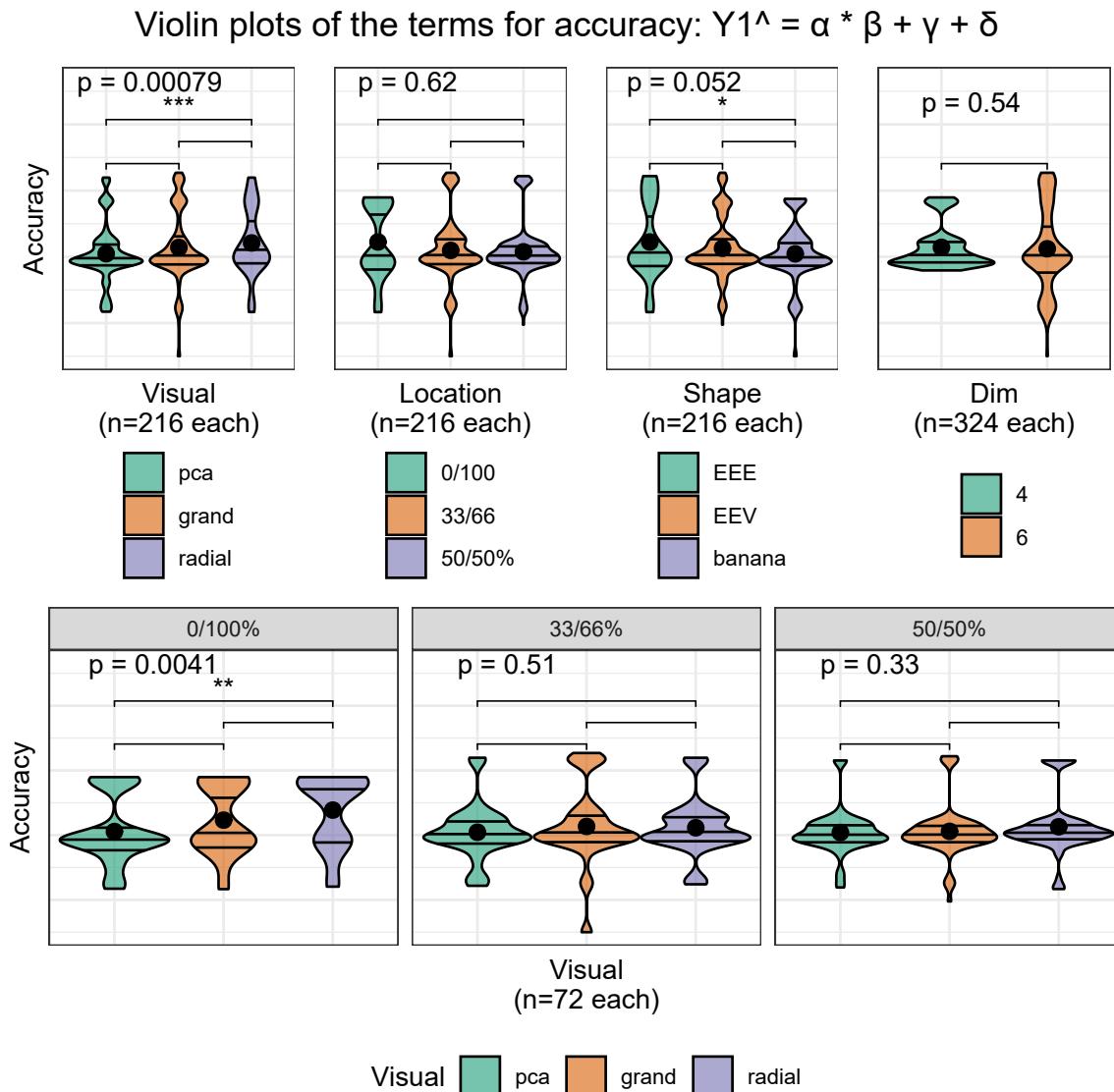
| Fixed effects | No. levels | No. terms | AIC        | BIC        | R2 cond.     | R2 marg.     | RMSE         |
|---------------|------------|-----------|------------|------------|--------------|--------------|--------------|
| a             | 1          | 3         | <b>-71</b> | <b>-44</b> | 0.303        | 0.018        | 0.194        |
| a+b+c+d       | 4          | 8         | -45        | 4          | 0.334        | 0.056        | 0.194        |
| a*b+c+d       | 5          | 12        | -26        | 41         | 0.338        | 0.064        | 0.193        |
| a*b*c+d       | 8          | 28        | 28         | 167        | <b>0.383</b> | 0.108        | 0.19         |
| a*b*c*d       | 15         | 54        | 105        | 360        | 0.37         | <b>0.222</b> | <b>0.185</b> |

**Table 4.2:** The task accuracy model coefficients for  $\hat{Y} = \alpha \times \beta + \gamma + \delta$ , with visual = pca, location = 0/100%, shape = EEE, and dim = 4 held as baselines. Visual being radial is the fixed term with the strongest evidence supporting the hypothesis. Interacting with the location term, there is evidence suggesting radial performs with minimal improvement for 33/66% location mixing.

|                             | Estimate | Std. Error | df    | t value | Pr(> t )  |
|-----------------------------|----------|------------|-------|---------|-----------|
| (Intercept)                 | 0.10     | 0.06       | 16.1  | 1.54    | 0.143     |
| <b>Visual</b>               |          |            |       |         |           |
| Visualgrand                 | 0.06     | 0.04       | 622.1 | 1.63    | 0.104     |
| Visualradial                | 0.14     | 0.04       | 617.0 | 3.77    | 0.000 *** |
| <b>Fixed effects</b>        |          |            |       |         |           |
| Location33/66%              | -0.02    | 0.07       | 19.9  | -0.29   | 0.777     |
| Location50/50%              | -0.04    | 0.07       | 20.0  | -0.66   | 0.514     |
| ShapeEEV                    | -0.05    | 0.06       | 11.8  | -0.82   | 0.427     |
| Shapebanana                 | -0.09    | 0.06       | 11.8  | -1.54   | 0.150     |
| Dim6                        | -0.01    | 0.05       | 11.8  | -0.23   | 0.824     |
| <b>Interactions</b>         |          |            |       |         |           |
| Visualgrand:Location33/66%  | -0.02    | 0.06       | 588.9 | -0.29   | 0.774     |
| Visualradial:Location33/66% | -0.12    | 0.06       | 586.5 | -2.13   | 0.033 *   |
| Visualgrand:Location50/50%  | -0.03    | 0.06       | 591.6 | -0.47   | 0.641     |
| Visualradial:Location50/50% | -0.06    | 0.06       | 576.3 | -1.16   | 0.248     |

Table 4.1 compares the model summaries across increasing complexity. The  $\alpha \times \beta + \gamma + \delta$  model to is selected to examine in more detail as it has relatively high condition  $R^2$  and not overly complex interacting terms. Table 4.2 looks at the coefficients for this model. There is strong evidence suggesting a relatively large increase in accuracy from the radial tour, though there is evidence that almost of increase is lost under 33/66% mixing.

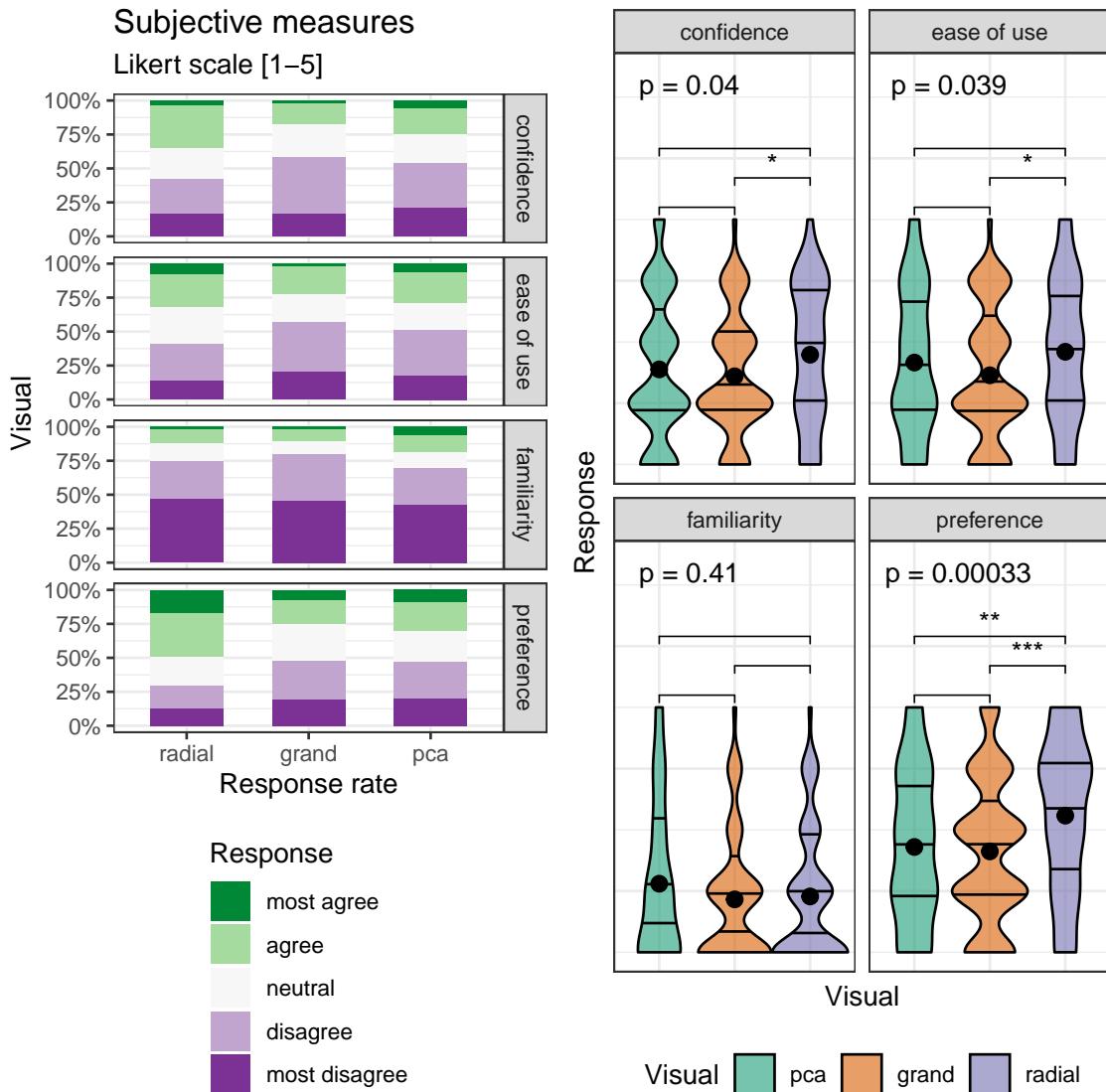
We also want to visually examine the conditional variables in the model. Figure 4.5 examines violin plots of accuracy by visual with panels distinguishing location (vertical) and shape (horizontal).



**Figure 4.5:** Violin plots of terms of the model  $\hat{Y} = \alpha \times \beta + \gamma + \delta$ . Overlaid with global significance from the Kruskal-Wallis test and pairwise significance from the Wilcoxon test, both are non-parametric, ranked-sum tests. Viewing marginal accuracy of the terms corroborate the primary findings that use of the radial tour leads to a significant increase in accuracy, at least over PCA, and this effect is particularly well supported when no location mixing is applied.

#### 4.2.2 Subjective measures

The 84 evaluations of the post-study survey also collect four subjective measures for each visual. Figure 4.6 shows the Likert plots, or stacked percentage bar plots, alongside violin plots with the same non-parametric, ranked sum tests previously used. Participants preferred to use radial for this task. Participants were also more confident of their answers and found radial tours easier than grand tours. All visuals have reportedly low familiarity, as expected from crowdsourced participants.



**Figure 4.6:** The subjective measures of the 84 responses of the post-study survey, five discrete Likert scale levels of agreement (L) Likert plots (stacked percent bar plots) with (R) violin plots of the same measures. Violin plots are overlaid with global significance from the Kruskal-Wallis test and pairwise significance from the Wilcoxon test. Both are non-parametric, ranked sum tests. Participants are more confident using the radial tour and find it easier to use than the grand tour. The radial tour is the most preferred visual.

### 4.3 Conclusion

Data visualization is an integral part of understanding relationships in data and how models are fitted. However, thorough exploration of data in high dimensions becomes difficult. Previous methods offer no means for an analyst to impact the projection basis. The manual tour provides a mechanism for changing the contribution of a selected variable to the basis. Giving analysts such control should facilitate the exploration of variable-level sensitivity to the identified structure.

This paper discussed a with-in participant user study ( $n = 108$ ) comparing the efficacy of three linear projection techniques: PCA, grand tour, and radial tour. The participants performed a supervised cluster task, explicitly identifying which variables contribute to the separation of two target clusters. This was evaluated evenly over four experimental factors. In summary, mixed model regression finds strong evidence that using the radial tour sizably increases accuracy, especially with the location of cluster separation is not mixed at 33/66%. The effect sizes on accuracy are large relative to the change from the other experimental factors and the random effect of data simulation, though smaller than the random effect of the participant. The radial tour was the most preferred of the three visuals.

There are several ways that this study could be extended. In addition to expanding the support of the experimental factors, more exciting directions include: introducing a new task, visualizations used, and experience level of the target population. It is difficult to achieve good coverage given the number of possible experimental factors.



## Chapter 5

# Exploring Local Explanations of Non-linear Models Using the Radial tour

The previous chapter discussed the within-participants user study comparing PCA, the grand tour, and the radial tour in a supervised variable attribution task. There was strong evidence that the radial tour led to a large increase in accuracy. The analyst can be more confident that the radial tour leads to a better analysis of variable-level attribution to variables identified in a projection.

Given the interpretability crisis of nonlinear models, it would be interesting to see if the radial tour can help. Specifically, this chapter investigates using the radial tour to explore variable sensitivity to the structure identified in linear local explanations of nonlinear models. That is, under what range of variable importance does an explanation make sense, and what contributions fail to support the prediction. This can be used to examine which variables lead to the misclassification of observation or an extreme residual. The radial tour can also test how susceptible a variable's contribution is to discerning the predictions of two observations.

The increased predictive power of nonlinear models comes at the cost of interpretability of its terms. This trade-off has led to the emergence of eXplainable AI (XAI). XAI attempts to shed light on how models use predictors to arrive at a prediction with *local explanations*, a point estimate of the linear variable importance in the vicinity of one observation. These

can be considered linear projections and can be further explored to understand better the interactions between variables used to make predictions across the predictive model surface. Here we describe interactive linear interpolation used for exploration at any observation and illustrate with examples with categorical (penguin species, chocolate types) and quantitative (soccer/football salaries, house prices) output. The methods are implemented in the **R** package **cheem**, available on CRAN.

Chapter 2 introduced predictive modeling, the interpretability crisis of nonlinear models, and local explanations — approximations of linear variable importance in the vicinity of one observation. The remainder of this chapter is organized as follows. Section 5.1 induces the SHAP and tree SHAP local explanation. Section 2.3.1 explains the animations of continuous linear projections. Section 5.2 discusses the visual layout in the graphical user interface, how it facilitates analysis, data preprocessing, and package infrastructure. Then Section 5.3 illustrates the application to supervised learning with categorical and quantitative output. Section 5.4 concludes with the insights gained and possible directions to explore in the future.

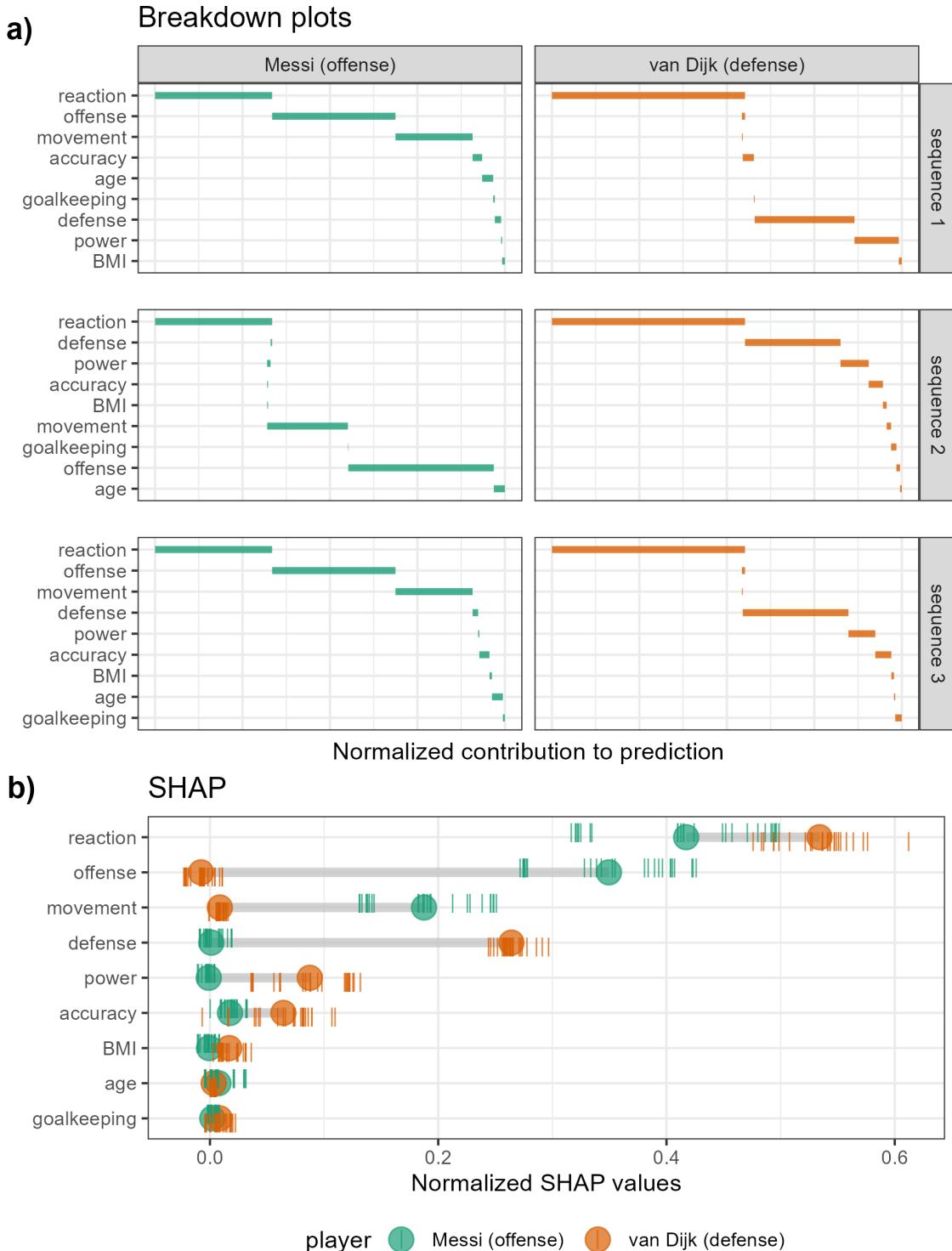
## 5.1 SHAP and tree SHAP local explanations

SHapley Additive exPlanations (SHAP) quantifies the variable contributions of one observation by examining the effect of other variables on the predictions. The explanations of SHAP refer to [Shapley \(1953\)](#)'s method to evaluate an individual's contribution to cooperative games by assessing this player's performance in the presence or absence of other players. [Strumbelj and Kononenko \(2010\)](#) introduced SHAP for local explanations in ML models. The attribution of variable importance depends on the sequence of the included variables. The SHAP values are the mean contributions over different variable sequences. The approach is related to partial dependence plots ([Molnar 2020](#)), used to explain the effect of a variable by predicting the response for a range of values on this variable after fixing the value of all other variables to their mean. Though partial dependence plots are a global approximation of the variable importance, while SHAP is specific to one observation. It could also be considered similar to examining the coefficients from all subsets regression,

as described in [Wickham, Cook, and Hofmann \(2015\)](#), helps to understand the relative importance of each variable in the context of all other candidate variables.

Following the use case *Explanatory Model Analysis* ([Biecek and Burzykowski 2021](#)), FIFA data is used to illustrate SHAP. Consider soccer data from the FIFA 2020 season ([Leone 2020](#)). There are 5000 observations of 9 skill measures (after aggregating highly correlated variables). A random forest model is fit regressing player's wages [2020 Euros] from their skill measurements. The SHAP values are compared for a star offensive player (L. Messi) and defensive player (V. van Dijk). The results are displayed in Figure 5.1. A difference in the attribution of the variable importance across the two positions of the players can be expected. This would be interpreted as how a player's salary depends on this combination of skill sets. Panel (a) is a modified breakdown plot ([Gosiewska and Biecek 2019](#)) where three sequences of variables are presented, so the two observations can be more easily compared. The magnitude of the contributions depends on the sequence in which they appear. Panel (b) shows the differences of the player's median values of over 25 such sequences. In summary, these plots highlight how local explanations bring interpretability to a model, at least in the vicinity of their observations. In this observation, two players with different positions receive different profiles of variable importance to explain the prediction of their wages.

For the application, we use *tree SHAP*, a variant of SHAP that enjoys a lower computational complexity ([Lundberg, Erion, and Lee 2018](#)). Instead of aggregating over sequences of the variables, tree SHAP calculates observation-level variable importance by exploring the structure of the decision trees. Tree SHAP is only compatible with tree-based models; random forests are used for illustration. The following section will use normalized SHAP values as a projection basis (call this the *attribution projection*) that will be used to explore the sensitivity of the variable contributions.



**Figure 5.1:** Illustration of SHAP values for a random forest model FIFA 2020 player wages from nine skill predictors. A star offensive and defensive player are compared, L. Messi and V. van Dijk, respectively. Panel (a) shows breakdown plots of three sequences of the variables. The sequence of the variables impacts the magnitude of their attribution. Panel (b) shows the distribution of attribution for each variable across 25 sequences of predictors, with the mean displayed as a dot for each player. Reaction skills are important for both players. Offense and movement are important for Messi but not van Dijk, and conversely, defense and power are important for van Dijk but not Messi.

## 5.2 The cheem viewer

To explore the local explanations, coordinated views ([Roberts 2007](#)) (also known as ensemble graphics, [Unwin and Valero-Mora 2018](#)) are provided in the *cheem viewer* application. There are two primary plots: the **global view** to give the context of all of the SHAP values and the **radial tour view** to explore the local explanations with user-controlled rotation. There are numerous user inputs, including variable selection for the radial tour and observation selection for making comparisons. There are different plots used for the categorical and quantitative responses. Figures [5.2](#) and [5.3](#) are screenshots showing the cheem viewer for the two primary tasks: classification (categorical response) and regression (quantitative response).

### 5.2.1 Global view

The global view provides a context of all observations and facilitates the exploration of the separability of the data- and attribution-spaces. Both of these spaces are of dimension  $n \times p$ , where  $n$  is the number of observations and  $p$  is the number of predictors. The attribution space corresponds to the local explanations for each observation, each a vector of  $p$  values.

Visualization is provided by the first two principal components of the data (left) and the attribution (middle) spaces. These single 2D projections will not reveal all of the structure of higher-dimensional space, but they are helpful visual summaries. In addition, a plot of the observed against predicted response values is also provided (Figures [5.2b](#), [5.3a](#)) to help identify observations poorly predicted by the model. For classification tasks, misclassified observations are circled in red. Linked brushing between the plots is provided, and a tabular display of selected points helps to facilitate exploration of the spaces and the model (shown in Figures [5.3d](#)).

While the comparison of these spaces is interesting, the primary purpose of the global view is to enable the selection of observations to explore the local explanations. The projection attribution of the primary observation (PO) is examined and typically viewed with an

optional comparison observation (CO). These observations are highlighted as asterisk and  $\times$ , respectively.

### 5.2.2 Radial tour

The local explanations for all observations are normalized (squared sum of values adds to 1), and thus, the relative importance of variables can be compared across all observations. These are depicted as vertical parallel coordinate plots ([Ocagne 1885](#)) on the basis biplot ([Gabriel 1971](#)). 1D biplot display depicts the values of the current basis as bars. The parallel coordinate overlays lines connecting one observation's variable attribution (Figures [5.2e](#) and [5.3e](#)). The attribution projections of the PO and CO are shown as dashed and dotted lines. From this plot, the range and density of the importance across all observations can be interpreted. For classification, one would look at differences between groups on any variable. For example, Figure [5.2e](#) suggests that `b1` is important for distinguishing the green class from the other two. For regression, one might generally observe which variables have low values for all observations (not important), for example, `BMI` and `pwr` in Figure [5.3e](#), and which have a range of high and low values (e.g. `off`, `def`), suggesting they are important for some observations and not important for other observations.

The overlaid bars on the parallel coordinate plot represent the attribution projection of the PO. (Remember that the PO is interactively selected from the global view). The attribution projection approximates the variable importance for predicting this observation. This combination of variables best explains the difference between the mean response and an observation's predicted value. It is not an indication of the local shape of the model surface. That is, it is not some indication of the tangent to the curve at this point.

The attribution projection of the PO is the initial 1D basis in a radial tour, displayed as a density plot for a categorical response (Figure [5.2f](#)) and as scatterplots for a quantitative response (Figure [5.3f](#)). The PO and CO are indicated by vertical dashed and dotted lines, respectively. The radial tour varies the contribution of the selected variable between 0-1. This is viewed as an animation of the projections from many intermediate bases. Doing so tests the sensitivity of structure (class separation or strength of relationship) to the variable's contribution. For classification, if the separation between classes diminishes

when the variable contribution is reduced, this suggests that the variable is important for class separation. For regression, if the relationship scatterplot weakens when the variable contribution is reduced, indicating that the variable is important for accurately predicting the response.

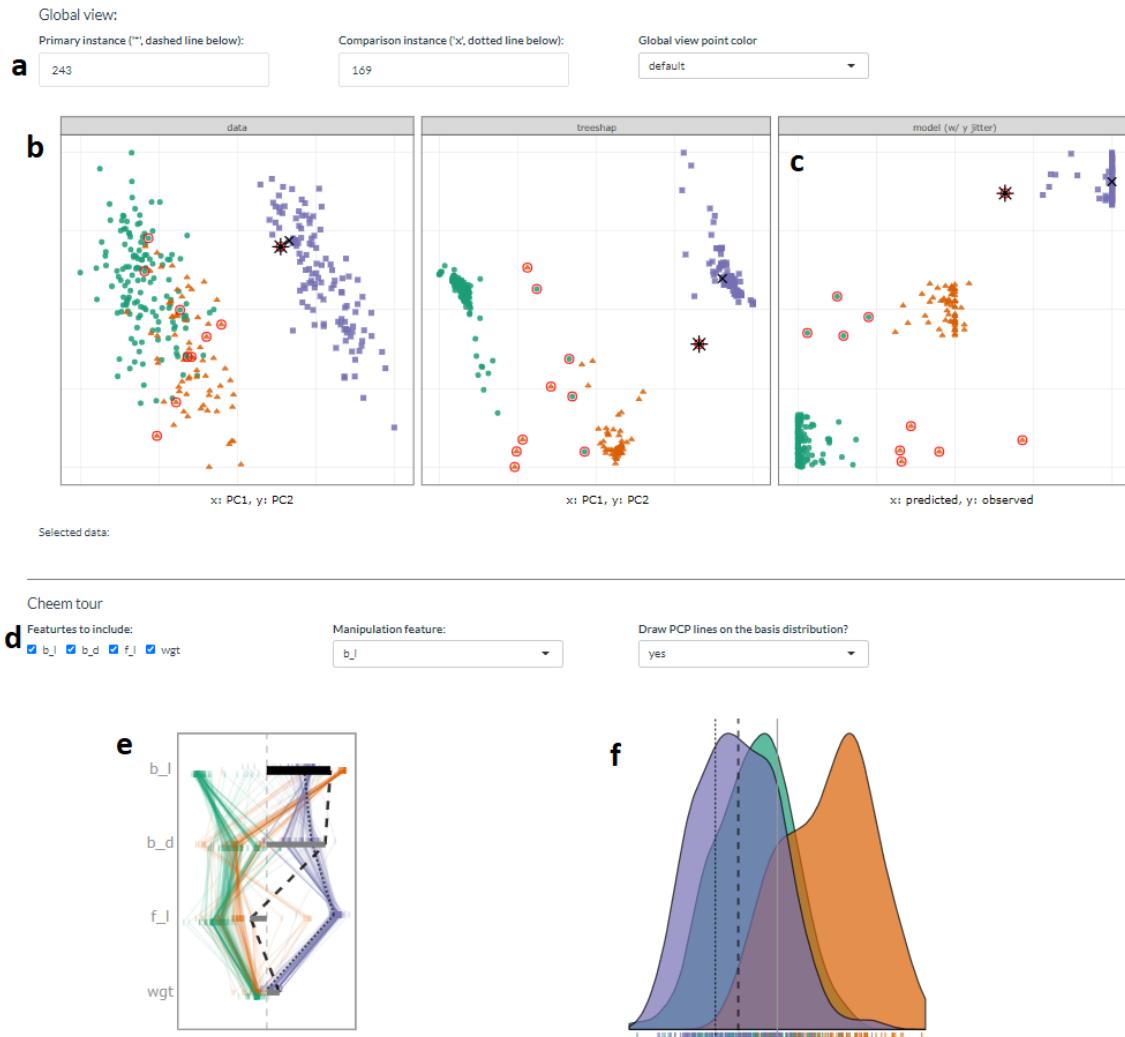
### 5.2.3 Classification task

Selecting a misclassified observation as PO and a correctly classified point nearby in data space as CO makes it easier to examine the variables most responsible for the error. The global view (Figure 5.2c) displays the model confusion matrix. The radial tour is 1D and displays as density where color indicates class. An animation slider enables the user to vary the contribution of each variable to explore the sensitivity of the separation to that variable.

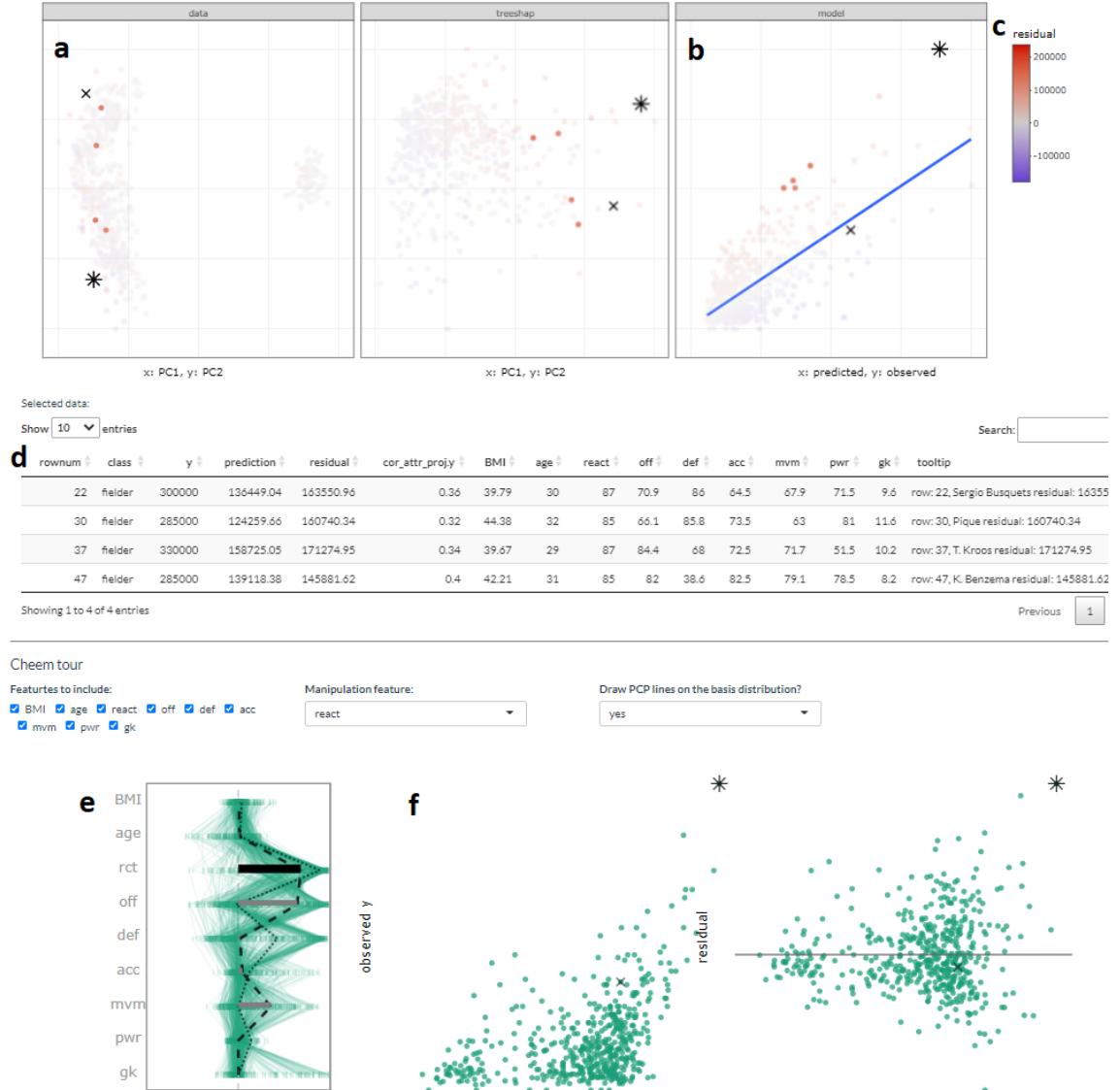
### 5.2.4 Regression task

Selecting an inaccurately predicted observation as PO and an accurately predicted observation with similar variable values as CO is a helpful way to understand how the model is failing or not. The global view (Figure 5.3a) shows a scatterplot of the observed vs predicted values, which should exhibit a strong relationship if the model is a good fit. The points can be colored by a statistic, residual, a measure of outlyingness (log Mahalanobis distance), or correlation aid in understanding the structure identified in these spaces.

In the radial tour view, the observed response and the residuals (vertical) are plotted against the attribution projection of the PO (horizontal). The attribution projection can be interpreted similarly to the predicted value from the global view plot. It represents a linear combination of the variables, and a good fit would be indicated when there is a strong relationship with the observed values. This can be viewed as a local linear approximation if the fitted model is nonlinear. As the contribution of a variable is varied, if the value of the PO does not change much, it would indicate that the prediction for this observation is NOT sensitive to that variable. Conversely, if the predicted value varies substantially, the prediction is very sensitive to that variable, suggesting that the variable is very important for the PO's prediction.



**Figure 5.2:** Overview of the cheem viewer for classification tasks. Global view inputs, (a), set the PO, CO, and color statistic. Global view, (b) PC1 by PC2 approximations of the data space and attribution space. (c) prediction by observed y (visual of the confusion matrix for classification tasks). Points are colored by predicted class, and red circles indicate misclassified observations. Radial tour inputs (d) select variables to include and which variable is changed in the tour. (e) shows a parallel coordinate display of the distribution of the variable attributions while bars depict contribution for the current basis. The black bar is the variable being changed in the radial tour. Panel (f) is the resulting data projection indicated as density in the classification case.



**Figure 5.3:** Overview of the cheem viewer for regression and illustration of interactive variables. Panel (a) PCA of the data and attributions spaces, (b) residual plot, predictions by observed values. Four selected points are highlighted in the PC spaces and tabularly displayed. Coloring on a statistic (c) highlights structure organized in the attribution space. Interactive tabular display (d) populates when observations are selected. Contribution of the 1D basis affecting the horizontal position (e) parallel coordinate display of the variable attribution from all observations, and horizontal bars show the contribution to the current basis. Regression projection (f) uses the same horizontal projection and fixes the vertical positions to the observed y and residuals (middle and right).

### 5.2.5 Interactive variables

The application has several reactive inputs that affect the data used, aesthetic display, and tour manipulation. These reactive inputs make the software flexible and extensible (Figure 5.2a & d). The application also has more exploratory interactions to help link points across displays, reveal structures found in different spaces, and access the original data.

A tooltip displays observation number/name and classification information while the cursor hovers over a point. Linked brushing allows the selection of points (left click and drag) where those points will be highlighted across plots (Figure 5.2a & b). The information corresponding to the selected points is populated on a dynamic table (Figure 5.2d). These interactions aid exploration of the spaces and, finally, the identification of primary and comparison observations.

### 5.2.6 Preprocessing

It is vital to mitigate the render time of visuals, especially when users may want to iterate many explorations. All computational operations should be prepared before runtime. The work remaining when an application is run solely reacts to inputs and rendering visuals and tables. Below discusses the steps and details of the reprocessing.

- **Data:** predictors and response are unscaled complete numerical matrix. Most models and local explanations are scale-invariant.
- **Model:** any model and compatible explanation could be explored with this method. Currently, random forest models are applied via the package **randomForest** (Liaw and Wiener 2002), compatibility tree SHAP. Modest hyperparameters are used, namely: 125 trees, number variables randomly sampled at each split,  $mtry = \sqrt{p}$  or  $p/3$  for classification and regression, and minimum size of terminal nodes  $\max(1, n/500)$  or  $\max(5, n/500)$  for classification and regression.
- **Local explanation:** Tree SHAP is calculated for *each* observation using the package **treeshap** (Kominsarczyk et al. 2021). This implementation aggregates exhaustively over all trees' attribution and does not fit interactions of variables.

- **Cheem viewer:** after the model and full explanation space are calculated, each variable is scaled by standard deviations away from the mean to achieve common support for visuals. Statistics for mapping to color are computed on the scaled spaces.

The time to preprocess the data will vary significantly with the complexity of the model and local explanation. For reference, the FIFA data contained 5000 observations of nine explanatory variables took 2.5 seconds to fit a random forest model of modest hyperparameters. Extracting the tree SHAP values of each observation took 270 seconds total. PCA and statistics of the variables and attributions took 2.8 seconds. These runtimes were from a non-parallelized R session on a modern laptop, but suffice to say that most of the time will be spent on the local attribution. An increase in model complexity or data dimensionality will quickly become an obstacle. Its reduced computational complexity makes tree SHAP an excellent candidate to start. (Alternatively, the package **fastshap** (Greenwell 2020) claims extremely low runtimes, attributed to fewer calls to the prediction function, partial implementation in C++, and efficient use of logical subsetting.)

### 5.2.7 Package infrastructure

The above-described method and application are implemented as an open-source **R** package, **cheem** available on [CRAN](#). Preprocessing was facilitated with models created via **randomForest** (Liaw and Wiener 2002) and explanations calculated with **treeshap** (Kominsarczyk et al. 2021). The application was made with **shiny** (W. Chang et al. 2021). The tour visual is built with **spinifex** (Spyrison and Cook 2020). Both views are created first with **ggplot2** (Wickham 2016) and then rendered as interactive **html** widgets with **plotly** (Sievert 2020). **DALEX** (Biecek 2018) and the free ebook, *Explanatory Model Analysis* (Biecek and Burzykowski 2021) were a boon to understanding local explanations and how to apply them.

### 5.2.8 Installation and getting started

The package can be installed from GitHub using the following **R** code:

```
install.packages("cheem", dependencies = TRUE)
library("cheem")
run_app()
```

Alternatively,

- a version of the cheem viewer shiny app can be directly accessed at [eb-smonash.shinyapps.io/cheem\\_initial/](https://eb-smonash.shinyapps.io/cheem_initial/).
- the development version of the package is available at <https://github.com/nspyrison/cheem>, and
- documentation of the package can be found at <https://nspyrison.github.io/cheem/>.

Follow the examples provided with the package to compute the local explainers (using `?cheem_ls`). The application expects the output returned by `cheem_ls()`, saved to an .rds file with `saveRDS()`, to be uploaded.

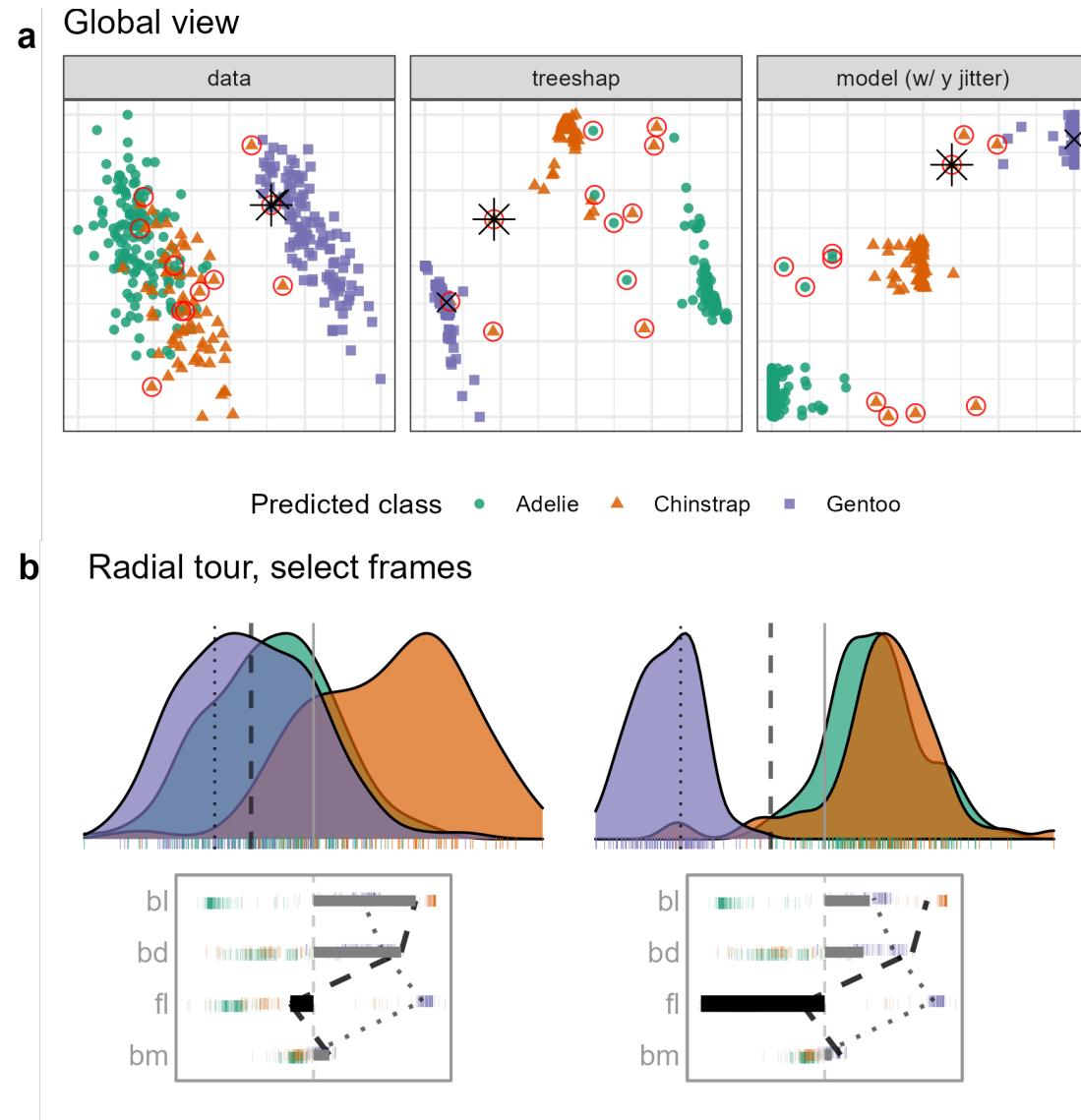
## 5.3 Case studies

To illustrate the cheem method, it is applied to modern datasets, two classification examples and then two of regression.

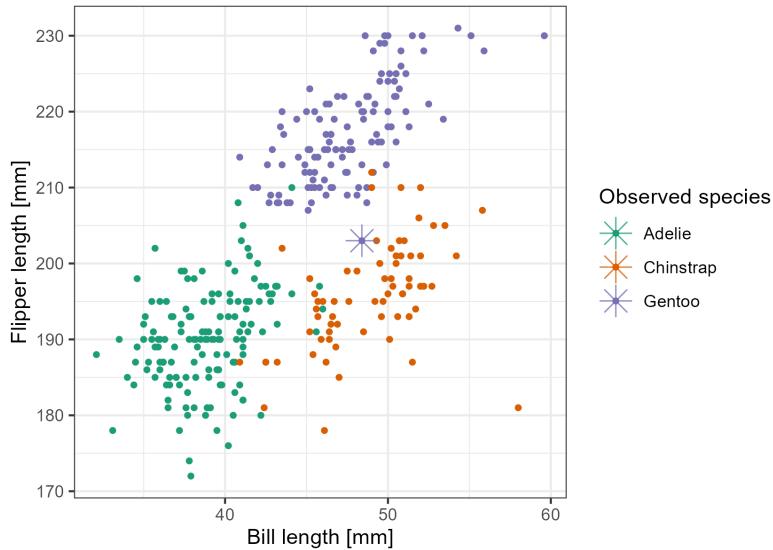
### 5.3.1 Palmer penguin, species classification

The Palmer penguins data ([Gorman, Williams, and Fraser 2014](#); [Horst, Hill, and Gorman 2020](#)) was collected on three species of penguins foraging near Palmer Station, Antarctica. The data was publicly available to substitute for the overly-used iris data and is quite similar in form. After removing incomplete observations, there are 333 observations of four physical measurements, bill length (`b1`), bill depth (`bd`), flipper length (`f1`), body mass (`bm`), for this illustration. A random forest model was fit with species as the response variable.

Figure 5.4 shows plots from the cheem viewer for exploring the random forest model on the penguins data. Panel (a) shows the global view, and panel (b) shows several 1D projections



**Figure 5.4:** Examining the SHAP values for a random forest model classifying Palmer penguin species. The PO is a Gentoo (purple) penguin that is misclassified as a Chinstrap (orange), marked as an asterisk in (a), and the dashed vertical line in (b). The radial view shows varying the contribution of *fl* from the initial attribution projection (b, left), which produces a linear combination where the PO is more probably (higher density value) a Chinstrap than a Gentoo (b, right). (The animation of the radial tour is at [vimeo.com/666431172](https://vimeo.com/666431172).)



**Figure 5.5:** Checking what is learned from the cheem viewer. This is a plot of flipper length ( $fl$ ) and bill length ( $bl$ ), where an asterisk highlights the PO. A Gentoo (purple) misclassified as a Chinstrap (orange). The PO has an unusually small  $fl$  length which is why it is confused with a Chinstrap.

generated with the radial tour. Penguin 243, a Gentoo (purple), is the PO because it has been misclassified as a Chinstrap (orange).

There is more separation visible in the attribution space than the data space, as would be expected. The predicted vs observed plot reveals a handful of misclassified observations. A Gentoo that has been wrongly labeled as a Chinstrap is selected for illustration. The PO is a misclassified point (represented by the asterisk in the global view and a dashed vertical line in the tour view). The CO is a correctly classified point (represented by an  $\times$  and a vertical dotted line).

The radial tour starts from the attribution projection of the misclassified observation (b, left). The important variables identified by SHAP in the (wrong) prediction for this observation are mostly  $bl$  and  $bd$  with small contributions of  $f1$  and  $bm$ . This projection is a view where the Gentoo (purple) looks much more likely for this observation than Chinstrap. That is, this combination of variables is not particularly useful because the PO looks very much like other Gentoo penguins. To explore this, the radial tour is used to vary the contribution of flipper length ( $f1$ ). (In our exploration, this was the third variable explored. It is typically helpful to explore the variables with more significant contributions, here  $bl$  and  $bd$ . Still, when doing this, nothing was revealed about how the PO differed

from other Gentoos). On varying  $f1$  as it contributes increasingly to the projection (b, right), more and more, this penguin looks like a Chinstrap. This suggests that  $f1$  should be considered an important variable for explaining the (wrong) prediction.

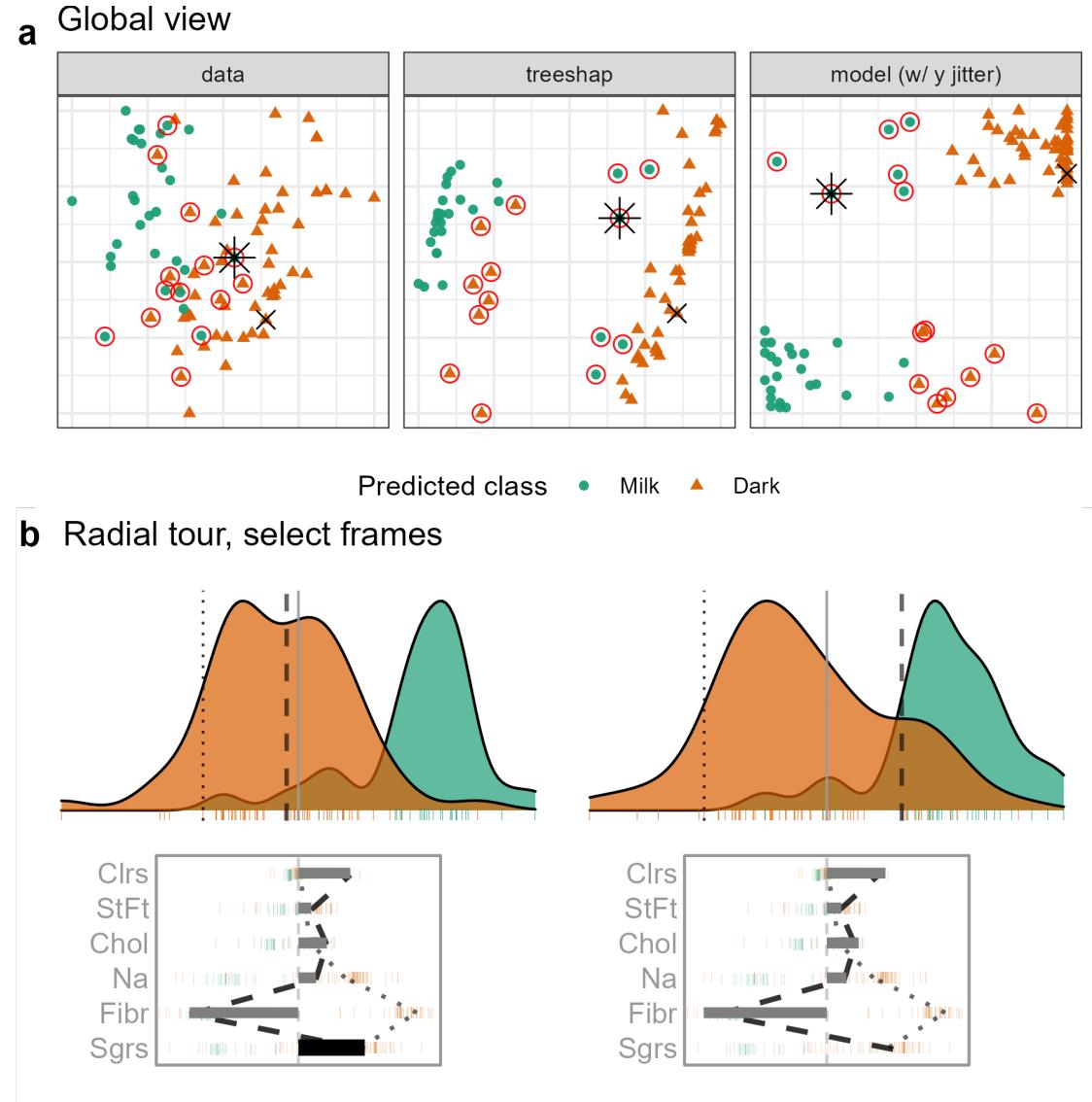
Figure 5.5 confirms that flipper length ( $f1$ ) is vital for the confusion of the PO as a Chinstrap. Here, flipper length and body length are plotted, and the PO can be seen to be closer to the Chinstrap group in these two variables, mainly because it has an unusually low value of flipper length relative to other Gentoos. From this view, it makes sense that it is a hard observation to account for as decision trees can only partition only vertical and horizontal lines.

### 5.3.2 Chocolates, milk/dark chocolate classification

The chocolates dataset consists of 88 observations of ten nutritional measurements determined from their labels and labeled as either milk or dark. Dark chocolate is considered healthier than milk. The data were collected by students during the Iowa State University class STAT503 from nutritional information from the manufacturer's website and normalized to 100g equivalents. The data is available in the **cheem** package. A random forest model is used for the classification of chocolate types.

It could be interesting to examine the nutritional properties of any dark chocolates that have been misclassified as milk. A reason to do this is that a dark chocolate, nutritionally more like milk, should not be considered a healthy alternative. It is interesting to explore which nutritional variables contribute most to misclassification.

This type of exploration is shown in Figure 5.6, where a chocolate labeled dark but predicted to be milk is chosen as the PO (observation 22). It is compared with a CO that is a correctly classified dark chocolate (observation 7). The PCA plot and the tree SHAP PCA plots (a) show a big difference between the two chocolate types but with confusion for a handful of observations. The misclassifications are more apparent in the observed vs predicted plot and can be seen to be mistaken in both ways: milk to dark and dark to milk.



**Figure 5.6:** Examining the local explanation for a PO which is dark (orange) chocolate incorrectly predicted to be milk (green). From the attribution projection, this chocolate correctly looks more like dark than milk, which suggests that the local explanation does not help understand the prediction for this observation. So, the contribution of Sugar is varied — reducing it corresponds primarily with increasing Fiber. When Sugar is zero, Fiber contributes strongly towards the left. In this particular view, the PO is closer to the bulk of the milk chocolates, suggesting that the prediction put a lot of importance on Fiber. This chocolate is a rare dark chocolate without any Fiber leading to it being mistaken for a milk chocolate. (A video of the tour animation can be found at [vimeo.com/666431143](https://vimeo.com/666431143).)

The attribution projection for chocolate 22 suggests that Fiber, Sugars, and Calories are most responsible for its incorrect prediction. The way to read this plot is to see that Fiber has a large negative value, while Sugars and Calories have reasonably large positive values. In the density plot, observations on the very left of the display would have high values of Fiber (matching the negative projection coefficient) and low values of Sugars and Calories. The opposite would be interpreting a point with high values in this plot. The dark chocolates (orange) are primarily on the left, and this is a reason why they are considered to be healthier: high fiber and low sugar. The density of milk chocolates is further to the right, indicating that they generally have low fiber and high sugar.

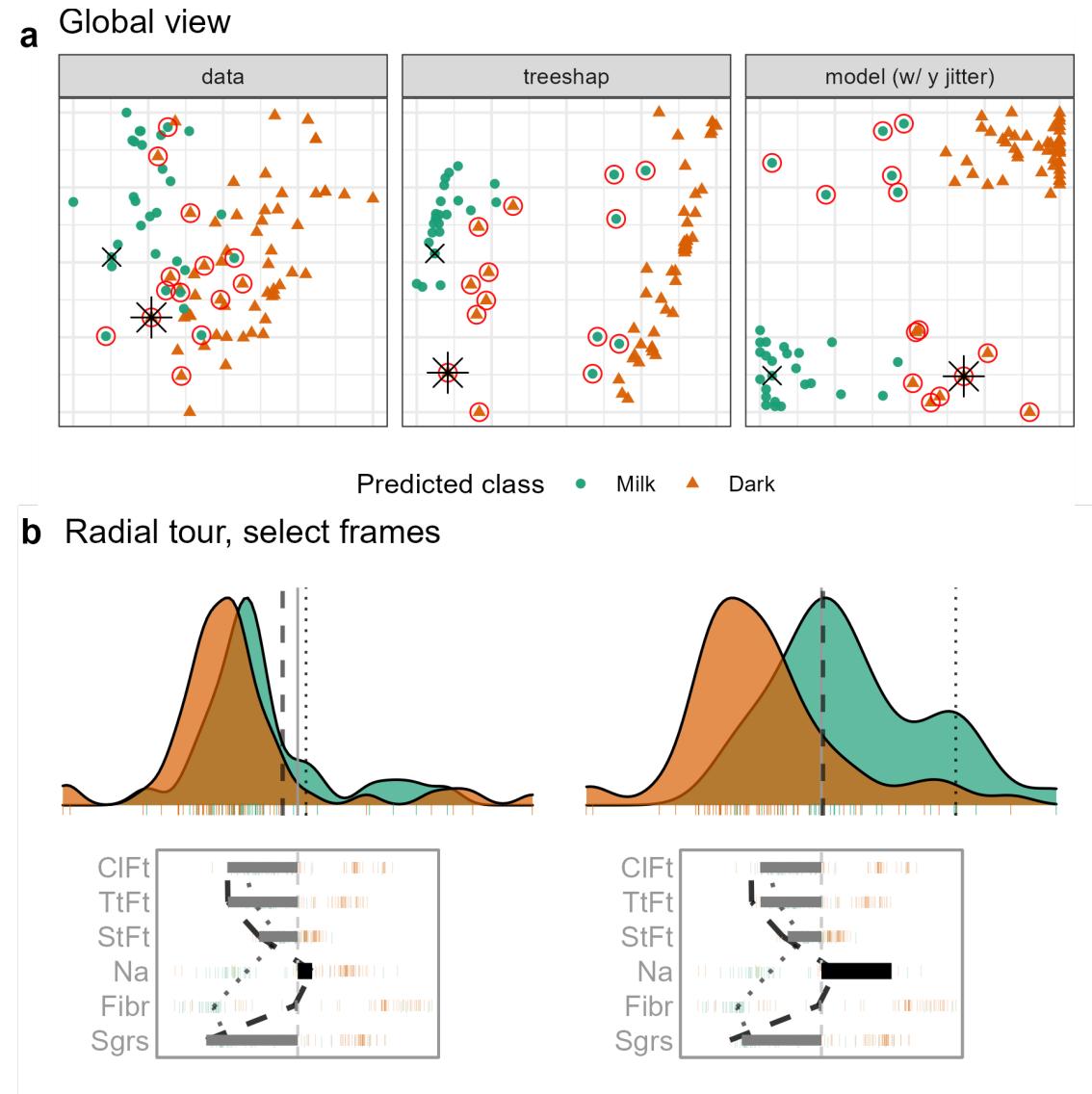
The PO (dashed line) can be viewed against the CO (dotted line). Now, one needs to pay attention to the parallel plot of the SHAP values, which are local to a particular observation, and the density plot, which is the same projection of all observations as specified by the SHAP values of the PO. The variable contributions to the two different predictions can be quickly compared in the parallel coordinate plot. The PO differs from the comparison primarily on the Fiber variable, which suggests that this is the reason for the incorrect prediction.

From the density plot, which is the attribution projection corresponding to the PO, both observations are more like dark chocolates. Varying the contribution of Sugars and altogether removing it from the projection is where the difference becomes apparent. When primarily Fiber is examined, observation 22 looks more like a milk chocolate.

It would also be interesting to explore an inverse misclassification. In this case, a milk chocolate is selected while it was misclassified as a dark chocolate. Chocolate 84 is just this case and is compared with a correctly predicted milk chocolate (observation 71). The corresponding global view and radial tour frames are shown in Figure 5.7.

The difference of position in the tree SHAP PCA with the previous case is quite significant; this gives a higher-level sense that the attributions should be quite different. Looking at the attribution projection, this is found to be the case. Previously, Fiber was essential while it is absent from the attribution in this case. Conversely, Calories from Fat and Total Fat have high attributions here, while they were unimportant in the preceding case.

---



**Figure 5.7:** Examining the local interpretation for a PO which is milk (green) chocolate incorrectly predicted to be dark (orange). In the attribution projection, the PO could be either milk or dark. Sodium and Fiber have the largest differences in attributed variable importance, with low values relative to other milk chocolates. The lack of importance attributed to these variables is suspected of contributing to the mistake, so the contribution of Sodium is varied. If Sodium had a larger contribution to the prediction (like in this view), the PO would look more like other milk chocolates. (A video of the tour animation can be found at [vimeo.com/666431148](https://vimeo.com/666431148).)

Comparing the attribution with the CO (dotted line), large discrepancies in Sodium and Fiber are identified. The contribution of Sodium is selected to be varied. Even in the initial projection, the observation looks slightly more like its observed milk than predicted dark chocolate. The misclassification appears least supported when the basis reaches sodium attribution of typical dark chocolate.

### 5.3.3 FIFA, wage regression

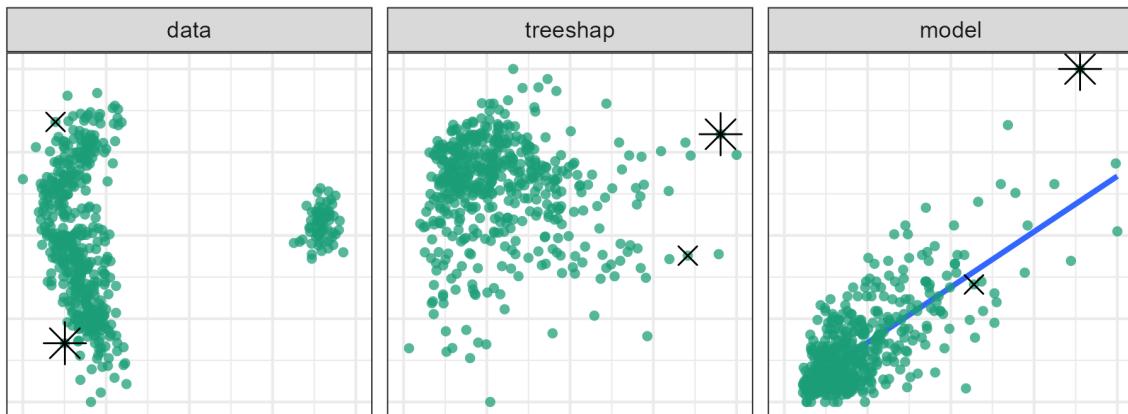
The 2020 season FIFA data ([Leone 2020](#); [Biecek 2018](#)) contains many skill measurements of soccer/football players and wage information. Nine higher-level skill groupings were identified and aggregated from highly correlated variables. A random forest model is fit from these predictors, regressing player wages [2020 euros]. The model was fit from 5000 observations before being thinned to 500 players to mitigate occlusion and render time. Continuing from the exploration in section [2.6](#), we are interested to see the difference in attribution based on the exogenous player position. That is, the model should be able to use multiple linear profiles to better predict the wages from different field positions of players despite not having this information. A leading offensive fielder (L. Messi) is compared with a top defensive fielder (V. van Dijk). The same observations were used in figure [5.1](#).

Figure [5.8](#), tests the support of the local explanation. Offensive and reaction skills (`off` and `rct`) are both crucial to explaining a star offensive player. If either of them were rotated out, the other would be rotated into the frame, maintaining a far-right position. However, Increasing the contribution of a variable with low importance would rotate both variables out of the frame.

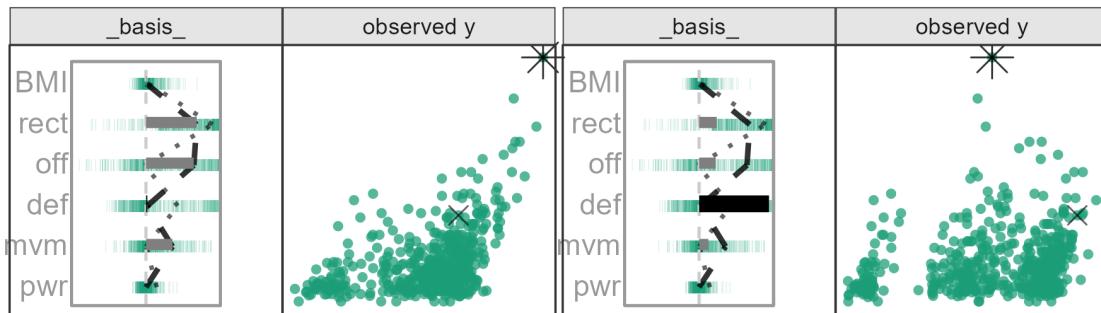
The contribution from `def` will be varied to contrast with offensive skills. As the contribution of defensive skills increases, Messi's is no longer separated from the group. Players with high values in defensive skills are now the rightmost points. In terms of what-if analysis, the difference between the data mean and his predicted wages would be halved if Messi's tree SHAP attributions at these levels.

**a**

Global view



**b** Radial tour, select frames



**Figure 5.8:** Exploring the wages (euros) relative to skill measurements in the FIFA 2020 data. Star offensive player (L. Messi) is the PO, and he is compared with a top defensive player (V. van Dijk). The attribution projection is shown at left, and it can be seen that this combination of variables produces a view where Messi has very high predicted (and observed) wages. Defense (`def`) is the chosen variable to vary. It starts off very low, and as its contribution is increased (right plot) it can be seen that Messi's predicted wages decrease dramatically. The increased contribution in defense comes at the expense of offensive and reaction skills. The interpretation is that, indeed, Messi's high wages are most attributable to his offensive and reaction skills, as initially provided by the local explanation. (A video of the animated radial tour can be found at [vimeo.com/666431163](https://vimeo.com/666431163).)

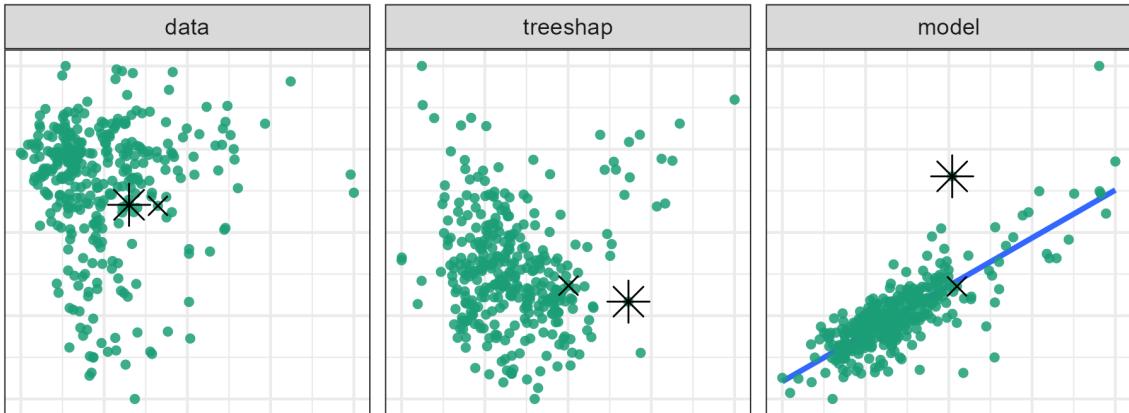
### 5.3.4 Ames housing 2018, sales price regression

Ames 2018, housing data was subset to North Ames (the neighborhood with the most house sales). The remaining are 338 house sales. A random forest model has fitted this price [USD] with the property variables: Lot Area (**LtA**), Overall Quality (**Qlt**), Year the house was Built (**YrB**), Living Area (**LvA**), number of Bathrooms (**Bth**), number of Bedrooms (**Bdr**), the total number of Rooms (**Rms**), Year the Garage was Built (**GYB**), and Garage Area (**GrA**). Using interaction from the global view, a house with an extreme negative residual and an accurate instance with a similar prediction are selected.

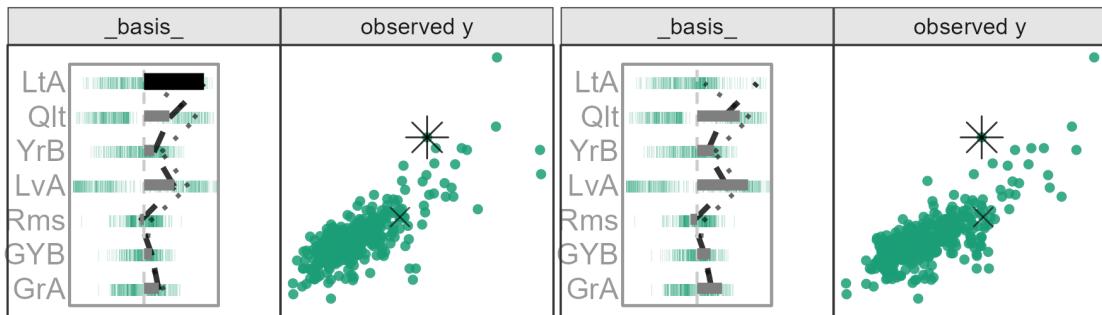
Figure 5.9 selects the house sale 74, a sizable under prediction with an enormous Lot Area contribution. The CO has a similar predicted price though the prediction was accurate and gives almost no attribution to lot size. The attribution projection places observations with high Living Areas to the right. The contribution of Living Area contrasts the contribution of this variable. As the contribution of Lot Area decreases, the predictive power decreases for the PO, while the CO remains stationary. This large of importance in the Living Area is relatively uncommon. Boosting tree models may be more resilient to such an under-prediction as they would up-weighting this residual and force its inclusion in the final model.

**a**

Global view



**b** Radial tour, select frames



**Figure 5.9:** Exploring an observation with a large residual as the PO from fitting sales price [USD] to other variables in the Ames housing 2018. The sale price of the PO was under-predicted. The local explanation indicates a sizable attribution to Lot Area (LtA). The CO has a similar predicted sales price and smaller residual and has minimal attribution to Lot Area. In the attribution projection, the PO has a higher sales price than the CO. Reducing the contribution of Lot Area brings these two prices in line. This suggests if the model did not consider Lot Area, then the two houses would be quite similar. That is, the large residual is due to a lack of factoring in the Lot Area for the prediction of PO's sales price. (A video showing the animation is at [vimeo.com/666431134](https://vimeo.com/666431134).)

## 5.4 Discussion

There is a clear need to extend the interpretability of black-box models. This chapter provides a technique that builds on local interpretations to explore the variable importance local to an observation. The local interpretations form an attribution projection from which variable contributions are varied using a radial tour. Several diagnostic plots are provided to assist with understanding the sensitivity of the prediction to particular variables. A global view shows the data space, explanation space, residual plot. The user can interactively select observations to compare, contrast, and study further. Then the radial tour is used to explore the variable sensitivity identified by the attribution projection.

This approach has been illustrated using four data examples of random forest models with the tree SHAP local explanation. In the penguins example, we showed how the misclassification of a penguin arose due to it having an unusually small flipper size compared to others of its species. This was verified by making a follow-up plot of the data. The chocolates example shows how a dark chocolate was misclassified primarily due to its attribution to Fiber, and a milk chocolate was misclassified as dark due to its lowish Sodium value. In the FIFA example, we show how low Messi's salary would be if it depended on defensive skills. In the Ames housing data, an inaccurate prediction for a house was likely due to the Lot Area not being effectively used by the random forest model.

This analysis is manually intensive and thus only feasible for investigating a few observations. The approach selects observations that the model has not done well to predict and compares it with an observation where it did fit well. The radial tour launches from the attribution projection to enable exploration of the sensitivity of the prediction to any variable. It can be helpful to make additional plots of the variables and responses to cross-check interpretations made from the cheem viewer. This methodology provides an additional tool in the box for studying model fitting.

An implementation is provided in the open-source **R** package **cheem**, available on [CRAN](#). Example data sets are provided, and you can upload your data after model fitting and computing the local explanations. In theory, this approach would work with any black-box

model, but the implementation currently only calculates tree SHAP for tree-based models supported by **treeshap** (tree based models from **gbm**, **lightgbm**, **randomForest**, **ranger**, or **xgboost** Greenwell et al. 2020; Shi et al. 2022; Liaw and Wiener 2002; Wright and Ziegler 2017; Chen et al. 2021, respectively). Tree SHAP was selected because of its computational efficiency. The SHAP and oscillation explanations could be added with the use of the **DALEX::explain()** and would be an excellent direction to extend the work (Biecek 2018; Biecek and Burzykowski 2021).

# Chapter 6

## Conclusion

Most data are multivariate. This thesis makes important contributions to the interactive visualization of quantitative multivariate data. Data visualization can convey complexities in data such as heteroskedasticity, irregular shapes of distributions, better than numerical summaries. Thus it is important activity for exploring data. However, it becomes more challenging when there are more than two variables (dimensions). It is common to tackle high-dimensions by reducing the dimensionality, and to visualize the lower dimensional data. This thesis provides new methodology using linear dimension reduction, built on dynamic animation of linear projections, with a specific approach called a radial tour. This has benefits over traditional approaches that use only a small handful of low-dimensional views, because many low-dimensional projections are shown. In particular, the radial tour is designed to help the user understand the relationship between a single variable and structure visible in the projections.

The Introduction lists the over-arching question: *Can the radial tour, with user-steering of the basis, help analysts' understanding of the variable sensitivity to structure in the projection?* This is divided into three research questions that were correspondingly addressed in the three content chapters. Chapter 3 developed a package addressing RQ 1 (How do we define and implement a user interface and interactions for the radial tours to add and remove variables smoothly from 1- and 2D linear data projections?). Chapter 4 discussed a user study to answer RQ 2; does the use of the interactive radial tour improve

analysts' understanding of the relationship between variables and structure in 2D linear projections compared to existing approaches? Chapter 5 covers the novel analysis and corresponding package responding to RQ 3, can the radial tours be used in conjunction with local explanations to improve the interpretability of black-box models?

## 6.1 Contributions

The contributions of this thesis can be split into scientific knowledge and software development.

### 6.1.1 Scientific knowledge

Chapter 3 clarifies the radial tour methodology, specifically using Rodrigues' rotation formula (Rodrigues 1840) to solve the rotation matrix defined for 1- and 2D tours. This use of the rotation formula sets up a scaffolding to extend the manual tour to three dimensions with another rotation angle to span the manipulation space. This work also supports radial tours by illustrating use cases on high-energy physics data sets.

Chapter 4 discusses a user study to evaluate the efficacy of the radial tour as compared with PCA and the grand tour. It defines a task and accuracy measure for a variable attribution of the separation of two clusters. The  $n = 108$  crowdsourced user study compares the performance of these visuals across three experimental factors of simulated data. Mixed model regression finds considerable evidence for a sizable improvement in accuracy from the radial tour, which participants also subjectively prefer.

Chapter 5 introduces cheem analysis. This extends the interpretability of nonlinear models by exploring local explanations with the radial tour. The global view gives a full observation summary of data space, attribution space, and residual plot side-by-side as a coordinated view. From this, an analyst identifies a primary observation to explore its explanation in detail and optionally compares it against another observation. The primary observation's normalized attribution becomes the starting basis for a radial tour. By varying the contributions of variables, an analyst tests the contribution's sensitivity to

the predictive power identified in the explanation. We provide usage and discussion from four contemporary datasets.

### 6.1.2 Software

The **R** package **spinifex** facilitates the creation of manual tours, which allow an analyst to steer the contributions of a variable. It creates a framework for the layered display of tours interoperable with the **tourr** package. This extension of the geometric display of tours will feel at home to **ggplot2** users. After composition, tours can be animated and exported as interactive or fixed animations. Vignettes and an interactive application help users rapidly understand the concepts facilitated by this work. The impact of **spinifex** can be seen in two ways. My contributions to **spinifex** and **tourr** won the ACEMS Impact and Engagement Award, 2018. Furthermore, the package is available on [CRAN](#) with vignettes and version notes on its [pkgdown](#) site. It has been downloaded over 15,800 times from CRAN between 09 April 2019 and 27 February 2022.

The **cheem** package facilitates the tree SHAP local explanations from tree-based models. Given a compatible model, functions perform the preprocessing to calculate the local explanations of all observations with several statistics that help describe the separability of data and attribution space. This processed object is then used to produce two novel visuals. The global view allows analysts to identify an observation to look at in more detail interactively. A custom display of the radial tour then helps evaluate the sensitivity of the variable contributions to the structure identified in the explanation. An interactive graphical user interface uses these visuals to facilitate the outlined cheem analysis. An interactive graphical user interface facilities this analysis. Several preprocessed datasets are included and allow analysts to upload their data after processing. This package was recently published to [CRAN](#) and has a corresponding [pkgdown](#) site.

## 6.2 Limitations

Below, we discuss several limitations with this work in chapter order. The following section echos this order and discusses possible directions to address these limitations.

---

Manual tours (and radial tours) are limited in several ways. For instance, the work currently lists rotation matrices for 1- and 2D manual tours. Another limitation is that they only change one variable at a time. This can be cumbersome and timely if many variables need to be zeroed or otherwise modified. The radial tour moves in three segments (total contribution, no contribution, initial contribution). It may be more approachable to directly relate the fraction of the slider to the magnitude of the variable's contribution.

Setting aside the manual tour, there are several extensions to the layered composition of tours in **spinifex**, such as extending the geometric display type. Because the manual tour is currently defined for  $d \in [1, 2]$  projections, the geometric display for  $d > 2$  **tourr**-made tours are under-supported. Most tour implementations are made with 2D monitors in mind. It would be nice to have implementations for extended reality with a stereoscopic head-mounted display.

The user study evaluating the radial tour has several intrinsic limitations from the choice of visuals and levels of the experimental factors. It only considered discrete 2D PCA, grand tour, and manual tour for supervised cluster separation on mostly linear clusters in four and six dimensions. Trials we collected online due to the ongoing COVID-19 pandemic.

The cheem analysis was illustrated using the tree SHAP explanations on random forest models. The **cheem** package handles several other tree-based models, though the analysis should be generalized to a broader base of models and compatible explanations. We also focus on quantitative matrices. This could be generalized to accommodate text, image, or time-series data. There is also no comparison evaluating the value of cheem analysis. It would be interesting to measure the benefit of cheem over local explanations or other observation-level evaluations of a model.

### 6.3 Future work

This section purposed possible directions to address or extend from the limitations outlined above.

Chapter 3 included the scaffolding to extend manual tours beyond 2D. Namely, this would require the use Rodrigues' rotation formula ([Rodrigues 1840](#)) to define another rotation

angle. The 3D projection would initialize a 4D manipulation space and use three input angles to control the contribution in the first three components. Another display dimension may benefit the detection and understanding of the higher dimensional structure, while we would also expect longer interaction time and less intuitive input.

In addition to the manual tour controlling the contribution of a single variable, it may be insightful to change the contributions of several variables at once (effectively manipulating a linear combination of variables). Varying combinations of variables may increase manipulation speed or prove to be unintuitive to input. Alternatively, an automated approach may help clean up variables with small contributions. A sort of dimension “reduction tour,” could append several manual tours that sequentially zero the contributions of variables contributing less than some threshold, which may prove to expedite analysis, especially for approaching a feature’s intrinsic dimensionality. This approach may be better as an initialization step before render time. This would abstract away some of the business and monotony of dealing with many variables allowing an analyst to focus on the variables that tangibly impact the projection.

Currently, the radial tour creates three segments (increase-decrease-increase of magnitude). An “ink tank” display for a manual tour may be more intuitive than the current approach. For this, the extent of contribution would match the progress of an animation slider. The first frame would contain zero variable contribution, and the last would have a total contribution. The starting frame could be the original contribution that would also be exaggerated or annotated to refer.

This last point is adjacent to the pedagogical side of tours and linear projections more broadly. The user-steering of the manual tour provides a way to play with projections and test hypotheses making it a good candidate as a learning tool. An interactive application built for teaching and exploring projection techniques would be a boon to scholars.

Extending the output dimensions of some tours is relatively easy to do. The display in **spinifex** has focused on  $d \in [1, 2]$ . Additional functions could be added to facilitate the geometric display of parallel coordinates, Chernoff faces, glyph-based, or pixel-based visuals.

For the reasons mentioned in Chapter 2, these displays are potentially best used with data with relatively few observations and many variables.

There are also several extensions to the display of 1 & 2D tours, including tabular displaying the numeric values of the basis, drawing of lines around convex and alpha hulls, and a *high-density region* displays with progressive density rugs, or a combination of point and density contour display where the bulk of the data is shown as density contour, while the outermost observations are displayed as points (Hyndman 1996; O’Hara-Wild et al. 2022).

It may be interesting to experience tours as 3D scatterplots in extended reality with stereoscopically accurate head-tracking may be fruitful. Nelson, Cook, and Cruz-Neira (1998) explore 2D tours in virtual reality. Other works view 3D scatterplot tours on 2D displays (Yang 1999, 2000). It would be interesting to see modern implementations using WebGL, Mozilla A-frame, or Unity. One concern would be keeping hardware and software as generalized as possible.

There are some ways that the user-study evaluating the radial tour could be extended. The data could be changed to involve nonlinear shapes, outliers, or varying density and dimensionality. Besides changing the support of the experimental factors to the user study, it may be more interesting to evaluate across different tasks or compare other visualization techniques — scatterplot matrices, parallel coordinate plots, or ridge plots of principal component approximations may be realistic. The performed user study crowdsourced participants with little exposure to linear projections. It would be interesting to compare the results from more experienced participants.

The outlined cheem analysis can be generally applied models and local explanations. While the package **cheem** currently calculates tree SHAP for tree-based models supported by **treeshap** (Kominsarczyk et al. 2021). This could be generalized more broadly to other models and local explanations. Those facilitated by DALEX::explain() seem to be an excellent direction to extend (Biecek 2018; Biecek and Burzykowski 2021). However, processing runtime is a looming obstacle that is exasperated when moving away from computationally efficient explanations. Alternatively, other statistics may better show the structure identified in the attribution space from the explanations.

---

In the cheem analysis, we had focused primarily on continuous numeric predictors. Perhaps this analysis could be extended to image, text, or time series analysis. The global view may remain helpful in summary and identification while observation-level exploration would likely need to change to fit the context of the data, be it saliency map, word-level contextual sentiment analysis, or other.

A user study would help elucidate the benefit of cheem analysis over local explanations or observation-specific analysis of a model. Our analysis can highlight the unique attribution of a selected observation against peers and test the sensitivity of that attribution. Perhaps a task identifying variable sensitivity to a prediction may be appropriate.



# Bibliography

- 10 Adadi, Amina, and Mohammed Berrada. 2018. “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI).” *IEEE Access* 6: 52138–60.
- Anderson, James A. 1995. *An Introduction to Neural Networks*. MIT press.
- Anscombe, F. J. 1973. “Graphs in Statistical Analysis.” *The American Statistician* 27 (1): 17–21. <https://doi.org/10.2307/2682899>.
- Arms, L., Dianne Cook, and Carolina Cruz-Neira. 1999. “The Benefits of Statistical Visualization in an Immersive Environment.” In *Virtual Reality, 1999. Proceedings.*, IEEE, 88–95. IEEE.
- Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, and Richard Benjamins. 2020. “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI.” *Information Fusion* 58: 82–115.
- Asimov, Daniel. 1985. “The Grand Tour: A Tool for Viewing Multidimensional Data.” *SIAM Journal on Scientific and Statistical Computing* 6 (1): 128–43. <https://doi.org/https://doi.org/10.1137/0906011>.
- Barrow, Madeleine, Jieyang Chong, and Nicholas Spyris. 2020. “Changes to Victorian Intraday Electricity Demand Following COVID-19 Restrictions.” In *Melbourne Datathon 2020, Insights Category*, 4. [https://github.com/nspyrison/melb\\_datathon2020/blob/master/\\_paper/paper.pdf](https://github.com/nspyrison/melb_datathon2020/blob/master/_paper/paper.pdf).

- Batch, Andrea, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. 2019. “There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics.” *IEEE Transactions on Visualization and Computer Graphics* 26 (1): 536–46.
- Becker, Richard A., and William S. Cleveland. 1987. “Brushing Scatterplots.” *Technometrics* 29 (2): 127–42. <https://doi.org/10.1080/00401706.1987.10488204>.
- Bertini, Enrico, Andrada Tatu, and Daniel Keim. 2011. “Quality Metrics in High-Dimensional Data Visualization: An Overview and Systematization.” *IEEE Transactions on Visualization and Computer Graphics* 17 (12): 2203–12.
- Besse, Philippe, Céline Castets-Renard, Aurélien Garivier, and Jean-Michel Loubes. 2019. “Can Everyday AI Be Ethical? Machine Learning Algorithm Fairness.” *Machine Learning Algorithm Fairness (May 20, 2018)*. *Statistiques Et Société* 6 (3).
- Biecek, Przemyslaw. 2018. “DALEX: Explainers for Complex Predictive Models in R.” *The Journal of Machine Learning Research* 19 (1): 3245–49.
- Biecek, Przemyslaw, and Tomasz Burzykowski. 2021. *Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models*. CRC Press.
- Box, George EP. 1976. “Science and Statistics.” *Journal of the American Statistical Association* 71 (356): 791–99.
- Breiman, Leo. 2001a. “Random Forests.” *Machine Learning* 45 (1): 5–32.
- . 2001b. “Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author).” *Statistical Science* 16 (3): 199–231.
- Brown, Timothy A. 2015. *Confirmatory Factor Analysis for Applied Research*. Guilford publications.
- Buja, Andreas, Dianne Cook, Daniel Asimov, and Catherine Hurley. 2005. “Computational Methods for High-Dimensional Rotations in Data Visualization.” In *Handbook of Statistics*, 24:391–413. Elsevier. [https://doi.org/10.1016/S0169-7161\(04\)24014-7](https://doi.org/10.1016/S0169-7161(04)24014-7).

- Cao, Nan, Yu-Ru Lin, David Gotz, and Fan Du. 2018. “Z-Glyph: Visualizing Outliers in Multivariate Data.” *Information Visualization* 17 (1): 22–40. <https://doi.org/10.1177/1473871616686635>.
- Card, Stuart K., Thomas P. Moran, and Allen Newell. 1983. *The Psychology of Human-Computer Interaction*. Crc Press.
- Cattell, Raymond B. 1966. “The Scree Test for the Number of Factors.” *Multivariate Behavioral Research* 1 (2): 245–76.
- Chambers, J. M., W. S. Cleveland, B. Kleiner, and P. A. Tukey. 1983. “Graphical Methods for Data Analysis.”
- Chang, Chunlei, Tim Dwyer, and Kim Marriott. 2018. “An Evaluation of Perceptually Complementary Views for Multivariate Data.” In *2018 IEEE Pacific Visualization Symposium (PacificVis)*, 195–204. IEEE.
- Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2021. *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. 2021. “Xgboost: Extreme Gradient Boosting.” <https://CRAN.R-project.org/package=xgboost>.
- Chernoff, Herman. 1973. “The Use of Faces to Represent Points in K-Dimensional Space Graphically.” *Journal of the American Statistical Association* 68 (342): 361–68. <https://doi.org/10.2307/2284077>.
- Cook, Dianne, and Andreas Buja. 1997. “Manual Controls for High-Dimensional Data Projections.” *Journal of Computational and Graphical Statistics* 6 (4): 464–80. <https://doi.org/10.2307/1390747>.
- Cook, Dianne, Andreas Buja, Javier Cabrera, and Catherine Hurley. 1995. “Grand Tour and Projection Pursuit.” *Journal of Computational and Graphical Statistics* 4 (3): 155. <https://doi.org/10.2307/1390844>.

- Cook, Dianne, Andreas Buja, Eun-Kyung Lee, and Hadley Wickham. 2008. “Grand Tours, Projection Pursuit Guided Tours, and Manual Controls.” In *Handbook of Data Visualization*, 295–314. Berlin, Heidelberg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-33037-0\\_13](https://doi.org/10.1007/978-3-540-33037-0_13).
- Cook, Dianne, Ursula Laa, and German Valencia. 2018. “Dynamical Projections for the Visualization of PDFSense Data.” *Eur. Phys. J. C* 78 (9): 742. <https://doi.org/10.1140/epjc/s10052-018-6205-2>.
- Dastin, Jeffrey. 2018. “Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women.” *Reuters*, October. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.
- Dimara, Evanthisia, and Charles Perin. 2019. “What Is Interaction for Data Visualization?” *IEEE Transactions on Visualization and Computer Graphics* 26 (1): 119–29.
- Díaz, Mark, Isaac Johnson, Amanda Lazar, Anne Marie Piper, and Darren Gergle. 2018. “Addressing Age-Related Bias in Sentiment Analysis.” In *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems*, 1–14.
- Duffy, Claire. 2019. “Apple Co-Founder Steve Wozniak Says Apple Card Discriminated Against His Wife.” *CNN*, November. <https://www.cnn.com/2019/11/10/business/goldman-sachs-apple-card-discrimination/index.html>.
- Espadoto, Mateus, Rafael M. Martins, Andreas Kerren, Nina S. T. Hirata, and Alexandru C. Telea. 2021. “Toward a Quantitative Survey of Dimension Reduction Techniques.” *IEEE Transactions on Visualization and Computer Graphics* 27 (3): 2153–73. <https://doi.org/10.1109/TVCG.2019.2944182>.
- Evans, Rhett, and Matthew Boreland. 2017. “Multivariate Data Analytics in PV Manufacturing—Four Case Studies Using Manufacturing Datasets.” *IEEE Journal of Photovoltaics* 8 (1): 38–47.
- Feurer, Matthias, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. 2015. “Efficient and Robust Automated Machine Learning.” *Advances in Neural Information Processing Systems* 28.

- Fisher, Ronald A. 1935. “The Logic of Inductive Inference.” *Journal of the Royal Statistical Society* 98 (1): 39–82.
- Friedman, Jerome H. 2002. “Stochastic Gradient Boosting.” *Computational Statistics & Data Analysis* 38 (4): 367–78.
- Gabriel, Karl Ruben. 1971. “The Biplot Graphic Display of Matrices with Application to Principal Component Analysis.” *Biometrika* 58 (3): 453–67.
- Galton, Francis. 1886. “Regression Towards Mediocrity in Hereditary Stature.” *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246–63.
- Gijsbers, Pieter, Florian Pfisterer, Jan N. van Rijn, Bernd Bischl, and Joaquin Vanschoren. 2021. “Meta-Learning for Symbolic Hyperparameter Defaults.” In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 151–52.
- Gorman, Kristen B., Tony D. Williams, and William R. Fraser. 2014. “Ecological Sexual Dimorphism and Environmental Variability Within a Community of Antarctic Penguins (Genus Pygoscelis).” *PloS One* 9 (3): e90081.
- Gosiewska, Alicja, and Przemyslaw Biecek. 2019. “IBreakDown: Uncertainty of Model Explanations for Non-Additive Predictive Models.” *arXiv Preprint arXiv:1903.11420*.
- Gracia, Antonio, Santiago González, Víctor Robles, Ernestina Menasalvas, and Tatiana von Landesberger. 2016. “New Insights into the Suitability of the Third Dimension for Visualizing Multivariate/Multidimensional Data: A Study Based on Loss of Quality Quantification.” *Information Visualization* 15 (1): 3–30. <https://doi.org/10.1177/1473871614556393>.
- Greenwell, Brandon. 2020. *Fastshap: Fast Approximate Shapley Values*. <https://CRAN.R-project.org/package=fastshap>.
- Greenwell, Brandon, Bradley Boehmke, Jay Cunningham, and G. B. M. Developers. 2020. “Gbm: Generalized Boosted Regression Models.” <https://CRAN.R-project.org/package=gbm>.

- Grinstein, Georges, Marjan Trutschl, and Urška Cvek. 2002. “High-Dimensional Visualizations,” 14.
- Hennerdal, Pontus. 2015. “Beyond the Periphery: Child and Adult Understanding of World Map Continuity.” *Annals of the Association of American Geographers* 105 (4): 773–90.
- Hevner, Alan R., Salvatore T. March, Jinsoo Park, and Sudha Ram. 2004. “Design Science in Information Systems Research.” *MIS Quarterly*, 75–105.
- Horst, Allison Marie, Alison Presmanes Hill, and Kristen B. Gorman. 2020. “Palmer-penguins: Palmer Archipelago (Antarctica) Penguin Data.” <https://allisonhorst.github.io/palmerpenguins/>.
- Huber, Wolfgang, Vincent J. Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S. Carvalho, Hector Corrada Bravo, Sean Davis, Laurent Gatto, and Thomas Girke. 2015. “Orchestrating High-Throughput Genomic Analysis with Bioconductor.” *Nature Methods* 12 (2): 115–21.
- Hutter, Frank, Lars Kotthoff, and Joaquin Vanschoren. 2019. *Automated Machine Learning: Methods, Systems, Challenges*. Springer Nature.
- Hyndman, Rob J. 1996. “Computing and Graphing Highest Density Regions.” *The American Statistician* 50 (2): 120–26.
- Kang, Yanfei, Rob J. Hyndman, and Kate Smith-Miles. 2017. “Visualising Forecasting Algorithm Performance Using Time Series Instance Spaces.” *International Journal of Forecasting* 33 (2): 345–58. <https://doi.org/10.1016/j.ijforecast.2016.09.004>.
- Keim, Daniel A. 2000. “Designing Pixel-Oriented Visualization Techniques: Theory and Applications.” *IEEE Transactions on Visualization and Computer Graphics* 6 (1): 59–78.
- Kleinberg, Jon, and Eva Tardos. 2006. *Algorithm Design*. Pearson Education India.
- Kodiyan, Akhil Alfons. 2019. “An Overview of Ethical Issues in Using AI Systems in Hiring with a Case Study of Amazon’s AI Based Hiring Tool.” *Researchgate Preprint*.

- Kominsarczyk, Konrad, Paweł Kozminski, Szymon Maksymiuk, and Przemysław Biecek. 2021. “Treeshap.” Model Oriented. <https://github.com/ModelOriented/treeshap>.
- Laa, Ursula, Dianne Cook, and German Valencia. 2020. “A Slice Tour for Finding Hollowness in High-Dimensional Data.” *Journal of Computational and Graphical Statistics* 29 (3): 681–87.
- Larson, Jeff, Surya Mattu, Lauren Kirchner, and Julia Angwin. 2016. “How We Analyzed the COMPAS Recidivism Algorithm.” *ProPublica*, May. [https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=RPR1E2qtzJltfJ0tS-gB\\_41kmfoWZAu4](https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=RPR1E2qtzJltfJ0tS-gB_41kmfoWZAu4).
- Lee, Stuart, Dianne Cook, Natalia da Silva, Ursula Laa, Nicholas Spyritson, Earo Wang, and H. Sherry Zhang. 2021. “The State-of-the-Art on Tours for Dynamic Visualization of High-Dimensional Data.” *WIREs Computational Statistics*, December, e1573. <https://doi.org/10.1002/wics.1573>.
- Lee, Stuart, Ursula Laa, and Dianne Cook. 2020. “Casting Multiple Shadows: High-Dimensional Interactive Data Visualisation with Tours and Embeddings.” *arXiv Preprint arXiv:2012.06077*.
- Leone, Stefano. 2020. “FIFA 20 Complete Player Dataset.” <https://kaggle.com/stefanoleone992/fifa-20-complete-player-dataset>.
- Lewis, Joshua, Laurens Van der Maaten, and Virginia de Sa. 2012. “A Behavioral Investigation of Dimensionality Reduction.” In *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 34.
- Liaw, Andy, and Matthew Wiener. 2002. “Classification and Regression by randomForest.” *R News* 2 (3): 18–22.
- Liu, Shusen, Dan Maljovec, Bei Wang, Peer-Timo Bremer, and Valerio Pascucci. 2017. “Visualizing High-Dimensional Data: Advances in the Past Decade.” *IEEE Transactions on Visualization and Computer Graphics* 23 (3): 1249–68. <https://doi.org/10.1109/TVCG.2016.2640960>.

- Lubischew, Alexander A. 1962. “On the Use of Discriminant Functions in Taxonomy.” *Biometrics*, 455–77. <https://doi.org/10.2307/2527894>.
- Lundberg, Scott M., Gabriel G. Erion, and Su-In Lee. 2018. “Consistent Individualized Feature Attribution for Tree Ensembles.” *arXiv Preprint arXiv:1802.03888*.
- Lundberg, Scott M., and Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions.” In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–77.
- Maaten, Laurens van der, Eric Postma, and Jaap Van den Herik. 2009. “Dimensionality Reduction: A Comparative Review.” *J Mach Learn Res* 10 (66-71): 13.
- Marriott, Kim, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas. 2018. *Immersive Analytics*. Springer.
- Matejka, Justin, and George Fitzmaurice. 2017. “Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics Through Simulated Annealing.” In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17*, 1290–94. Denver, Colorado, USA: ACM Press. <https://doi.org/10.1145/3025453.3025912>.
- Molnar, Christoph. 2020. *Interpretable Machine Learning*. Lulu. com. [christophm.github.io/interpretable-ml-book/](http://christophm.github.io/interpretable-ml-book/).
- Munzner, Tamara. 2014. *Visualization Analysis and Design*. AK Peters/CRC Press.
- Nelson, Laura, Dianne Cook, and Carolina Cruz-Neira. 1998. “XGobi Vs the C2: Results of an Experiment Comparing Data Visualization in a 3-D Immersive Virtual Reality Environment with a 2-D Workstation Display.” *Computational Statistics* 14 (1): 39–52.
- Nonato, Luis Gustavo, and Michael Aupetit. 2018. “Multidimensional Projection for Visual Analytics: Linking Techniques with Distortions, Tasks, and Layout Enrichment.” *IEEE Transactions on Visualization and Computer Graphics* 25 (8): 2650–73.

- O’Hara-Wild, Mitchell, Stephen Pearce, Ryo Nakagawara, Sayani Gupta, Darya Vanichkina, Emi Tanaka, Thomas Fung, and Rob Hyndman. 2022. “Gghdr: Visualisation of Highest Density Regions in ‘Ggplot2’” <https://CRAN.R-project.org/package=gghdr>.
- O’Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Ocagne, Maurice d’. 1885. *Coordonne’s Paralle’les Et Axiales. Me’thode de Transformation Ge’ome’trique Et Proce’dé Nouveau de Calcul Graphique de’duits de La Considé’ration Des Coordonne’s Paralle’les*. Paris: Gauthier-Villars.
- Palan, Stefan, and Christian Schitter. 2018. “Prolific. A Subject Pool for Online Experiments.” *Journal of Behavioral and Experimental Finance* 17: 22–27.
- Panagiotelis, Anastasios. 2020. “Manifold Learning.” Monash University.
- Pearson, Karl. 1901. “LIII. On Lines and Planes of Closest Fit to Systems of Points in Space.” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2 (11): 559–72.
- Pedersen, Thomas Lin, and David Robinson. 2020. “Gganimate: A Grammar of Animated Graphics.” <https://CRAN.R-project.org/package=gganimate>.
- Pfisterer, Florian, Jan N. van Rijn, Philipp Probst, Andreas C. Müller, and Bernd Bischl. 2021. “Learning Multiple Defaults for Machine Learning Algorithms.” In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 241–42.
- Probst, Philipp, Anne-Laure Boulesteix, and Bernd Bischl. 2019. “Tunability: Importance of Hyperparameters of Machine Learning Algorithms.” *The Journal of Machine Learning Research* 20 (1): 1934–65.
- Probst, Philipp, Marvin N. Wright, and Anne-Laure Boulesteix. 2019. “Hyperparameters and Tuning Strategies for Random Forest.” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9 (3): e1301.
- Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. 2016. “"Why Should I Trust You?": Explaining the Predictions of Any Classifier.” In *Proceedings of the 22nd*

- ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–44. KDD ’16. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939778>.
- Roberts, Jonathan C. 2007. “State of the Art: Coordinated & Multiple Views in Exploratory Visualization.” In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, 61–71. IEEE.
- Rodrigues, Olinde. 1840. “Des Lois géométriques Qui régissent Les déplacements d’un Système Solide Dans l’espace: Et de La Variation Des Cordonnées Provenant de Ces déplacements Considérés Indépendamment Des Causes Qui Peuvent Les Produire.” *Journal de Mathématiques Pures Et Appliquées* 5: 380–440.
- Scrucca, Luca, Michael Fop, T. Brendan Murphy, and Adrian E. Raftery. 2016. “Mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models.” *The R Journal* 8 (1): 289–317. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5096736/>.
- Sedlmair, Michael, Tamara Munzner, and Melanie Tory. 2013. “Empirical Guidance on Scatterplot and Dimension Reduction Technique Choices.” *IEEE Transactions on Visualization & Computer Graphics*, no. 12: 2634–43.
- Shapley, Lloyd S. 1953. *A Value for n-Person Games*. Princeton University Press.
- Shi, Yu, Guolin Ke, Damien Soukhavong, James Lamb, Qi Meng, Thomas Finley, Taifeng Wang, et al. 2022. “Lightgbm: Light Gradient Boosting Machine.” <https://CRAN.R-project.org/package=lightgbm>.
- Shmueli, Galit. 2010. “To Explain or to Predict?” *Statistical Science* 25 (3): 289–310.
- Shrikumar, Avanti, Peyton Greenside, and Anshul Kundaje. 2017. “Learning Important Features Through Propagating Activation Differences.” In *International Conference on Machine Learning*, 3145–53. PMLR.
- Shrikumar, Avanti, Peyton Greenside, Anna Shcherbina, and Anshul Kundaje. 2016. “Not Just a Black Box: Learning Important Features Through Propagating Activation Differences.” *arXiv Preprint arXiv:1605.01713*.

- Sievert, Carson. 2020. *Interactive Web-Based Data Visualization with r, Plotly, and Shiny*. Chapman; Hall/CRC. <https://plotly-r.com>.
- Simonyan, Karen, Andrea Vedaldi, and Andrew Zisserman. 2014. “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.” In *In Workshop at International Conference on Learning Representations*. Citeseer.
- Snyder, John Parr. 1987. *Map Projections–A Working Manual*. Vol. 1395. US Government Printing Office.
- Spyrison, Nicholas, and Dianne Cook. 2020. “Spinifex: An R Package for Creating a Manual Tour of Low-Dimensional Projections of Multivariate Data.” *The R Journal* 12 (1): 243. <https://doi.org/10.32614/RJ-2020-027>.
- Spyrison, Nicholas, Benjamin Lee, and Lonni Besançon. 2021. “"Is IEEE VIS \*That\* Good?" On Key Factors in the Initial Assessment of Manuscript and Venue Quality.” *OSF Preprints*, July. <https://doi.org/10.31219/osf.io/65wm7>.
- Strumbelj, Erik, and Igor Kononenko. 2010. “An Efficient Explanation of Individual Classifications Using Game Theory.” *The Journal of Machine Learning Research* 11: 1–18.
- Swayne, Deborah F, Duncan Temple Lang, Andreas Buja, and Dianne Cook. 2003. “GGobi: Evolving from XGobi into an Extensible Framework for Interactive Data Visualization.” *Computational Statistics & Data Analysis*, Data Visualization, 43 (4): 423–44. [https://doi.org/10.1016/S0167-9473\(02\)00286-4](https://doi.org/10.1016/S0167-9473(02)00286-4).
- Tukey, John W. 1977. *Exploratory Data Analysis*. Vol. 32. Pearson.
- Unwin, Antony, and Pedro Valero-Mora. 2018. “Ensemble Graphics.” *Journal of Computational and Graphical Statistics* 27 (1): 157–65. <https://doi.org/10.1080/10618600.2017.1383264>.
- Vanni, Laurent, Mélanie Ducoffe, Carlos Aguilar, Frédéric Precioso, and Damon Mayaffre. 2018. “Textual Deconvolution Saliency (TDS): A Deep Tool Box for Linguistic Analysis.” In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 548–57.

- Venna, Jarkko, and Samuel Kaski. 2006. “Visualizing Gene Interaction Graphs with Local Multidimensional Scaling.” In *ESANN*, 6:557–62. Citeseer.
- Wagner Filho, Jorge, Marina Rey, Carla Freitas, and Luciana Nedel. 2018. “Immersive Visualization of Abstract Information: An Evaluation on Dimensionally-Reduced Data Scatterplots.” In.
- Wang, Bo-Ting, T. J. Hobbs, Sean Doyle, Jun Gao, Tie-Jiun Hou, Pavel M. Nadolsky, and Fredrick I. Olness. 2018. “Mapping the Sensitivity of Hadronic Experiments to Nucleon Structure.” *Physical Review D* 98 (9): 094030. <https://doi.org/10.1103/PhysRevD.98.094030>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Dianne Cook, and Heike Hofmann. 2015. “Visualizing Statistical Models: Removing the Blindfold: Visualizing Statistical Models.” *Statistical Analysis and Data Mining: The ASA Data Science Journal* 8 (4): 203–25. <https://doi.org/10.1002/sam.11271>.
- Wickham, Hadley, Dianne Cook, Heike Hofmann, and Andreas Buja. 2011. “Tourr: An r Package for Exploring Multivariate Data with Projections.” *Journal of Statistical Software* 40 (2). <https://doi.org/10.18637/jss.v040.i02>.
- Wickham, Hadley, and Garrett Grolemund. 2017. *R for Data Science*. O’Reilly Media. <https://r4ds.had.co.nz/index.html>.
- Winer, Ben James. 1962. “Statistical Principles in Experimental Design.”
- Wright, Marvin N., and Andreas Ziegler. 2017. “Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software* 77 (1): 1–17. <https://doi.org/10.18637/jss.v077.i01>.
- Xie, Yihui. 2016. *Bookdown: Authoring Books and Technical Documents with R Markdown*. Boca Raton, Florida: Chapman; Hall/CRC. <https://github.com/rstudio/bookdown>.

- Yang, Li. 1999. “3d Grand Tour for Multidimensional Data and Clusters.” In *Advances in Intelligent Data Analysis*, edited by David J. Hand, Joost N. Kok, and Michael R. Berthold, 173–84. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. [https://doi.org/10.1007/3-540-48412-4\\_15](https://doi.org/10.1007/3-540-48412-4_15).
- . 2000. “Interactive Exploration of Very Large Relational Datasets Through 3d Dynamic Projections.” In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '00*, 236–43. Boston, Massachusetts, United States: ACM Press. <https://doi.org/10.1145/347090.347134>.
- Yao, Quanming, Mengshuo Wang, Yuqiang Chen, Wenyuan Dai, Yu-Feng Li, Wei-Wei Tu, Qiang Yang, and Yang Yu. 2019. “Taking Human Out of Learning Applications: A Survey on Automated Machine Learning.” *arXiv:1810.13306 [Cs, Stat]*, December. <http://arxiv.org/abs/1810.13306>.



# **Appendix A**

## **Glossary**

This thesis was written with statistics terminology. This glossary helps bridge the language used here to be more accessible to machine learning audiences.

---

| Term                  | Alias  |
|-----------------------|--|
| observation           | instance, item, case, row (of data)                  |
| variable              | feature, column (of data)                            |
| explanatory variables | independent-/input- variable, predictors, covariates |
| response variable     | predicted-/dependent-/target-/output- variable       |

---



## Appendix B

# Supplementary material

Below collects links of animations related to the content as well as vignettes and related documentation. The content for chapter 4 continues with a bit of an extended analysis looking at the visual displays in the user study, the participant demographics, and a parallel analysis regressing of log time.

### B.1 Thesis

| Content                          | Link  |
|----------------------------------|---|
| thesis repository                | <a href="https://shorturl.at/clqvB">https://shorturl.at/clqvB</a>     |
| thesis, pdf format               | <a href="https://shorturl.at/uBYZ5">https://shorturl.at/uBYZ5</a>     |
| thesis, html format              | <a href="https://shorturl.at/dftCN">https://shorturl.at/dftCN</a>     |
| Penguins radial tour, Figure 1.2 | <a href="https://vimeo.com/676723431">https://vimeo.com/676723431</a> |
| Penguins grand tour, Figure 2.4  | <a href="https://vimeo.com/676723441">https://vimeo.com/676723441</a> |

### B.2 Chapter 3, spinifex links

| Content                        | Link  |
|--------------------------------|---|
| spinifex documentation website | <a href="https://nspyrison.github.io/spinifex/">https://nspyrison.github.io/spinifex/</a> |

|  |   |
|--|---|
| spinifex vignette, Getting started with spinifex | <a href="https://shorturl.at/dgBOT">https://shorturl.at/dgBOT</a> |
| spinifex vignette, Ggproto api                   | <a href="https://shorturl.at/mwBP6">https://shorturl.at/mwBP6</a> |
| Fleas, radial tour animation, Figure             |   |
| <a href="#">3.3</a>                              | <a href="https://shorturl.at/yQT38">https://shorturl.at/yQT38</a> |
| Jet cluster, radial tour on PC1                  | <a href="https://shorturl.at/qIXZ8">https://shorturl.at/qIXZ8</a> |
| Jet cluster, radial tour on PC2                  | <a href="https://shorturl.at/ertFV">https://shorturl.at/ertFV</a> |
| Jet cluster, radial tour on PC3, Figure          |   |
| <a href="#">3.5 – least sensitive</a>            | <a href="https://shorturl.at/ehX68">https://shorturl.at/ehX68</a> |
| Jet cluster, radial tour on PC4, Figure          |   |
| <a href="#">3.4 – most senstive</a>              | <a href="https://shorturl.at/bcwN5">https://shorturl.at/bcwN5</a> |
| DIS cluster, radial tour on PC1                  | <a href="https://shorturl.at/cdqN4">https://shorturl.at/cdqN4</a> |
| DIS cluster, radial tour on PC2, Figure          |   |
| <a href="#">3.7 – least sensitive</a>            | <a href="https://shorturl.at/yzT13">https://shorturl.at/yzT13</a> |
| DIS cluster, radial tour on PC3                  | <a href="https://shorturl.at/kvwS8">https://shorturl.at/kvwS8</a> |
| DIS cluster, radial tour on PC4                  | <a href="https://shorturl.at/dowHN">https://shorturl.at/dowHN</a> |
| DIS cluster, radial tour on PC5                  | <a href="https://shorturl.at/insHR">https://shorturl.at/insHR</a> |
| DIS cluster, radial tour on PC6, Figure          |   |
| <a href="#">3.6 – most senstive</a>              | <a href="https://shorturl.at/ckGHS">https://shorturl.at/ckGHS</a> |

---

## B.3 Chapter 4, user study extended analysis

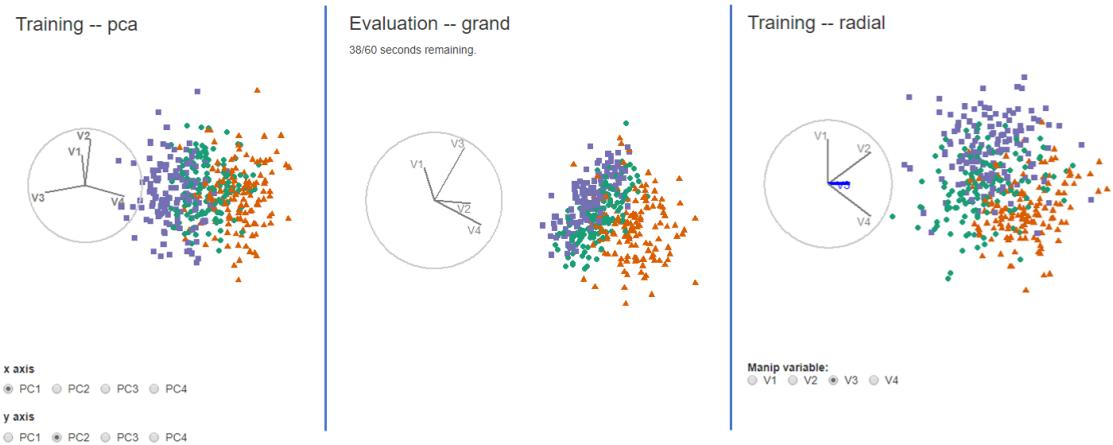
This section covers extended analysis. First, an illustrations of the different visuals are provided. Then, the demographics of the participants is covered. Lastly, a parallel modeling analysis on log response time is conducted.

### B.3.1 Illustrations of visual methods

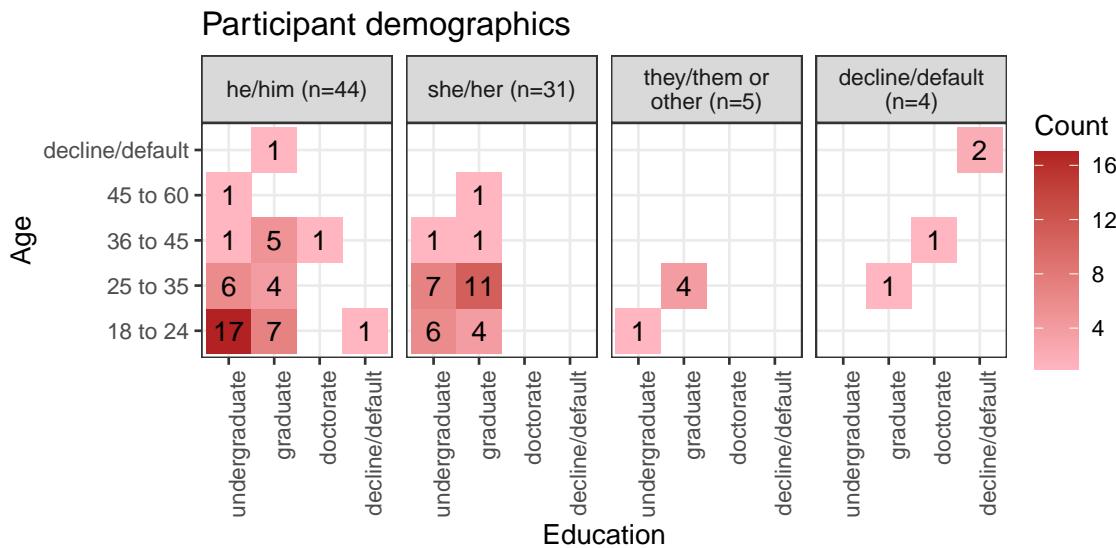
Below illustrates the three visual methods evaluated in the user study. Data was collected from a **shiny** application, and used pre-rendered **gif** files based on the selected inputs.

### B.3.2 Survey participant demographics

The target population is relatively well-educated people, as linear projections may prove difficult for generalized consumption. Hence Prolific.co participants are restricted to those with an undergraduate degree (58,700 of the 150,400 users at the study time). From this



**Figure B.1:** Examples of application for PCA, grand tour, and radial tour.



**Figure B.2:** Heatmaps of survey participant demographics; counts of age group by completed education as faceted across preferred pronoun. Our sample tended to be between 18 and 35 years of age with an undergraduate or graduate degree.

cohort, 108 performed a complete study. Of these participants, 84 submitted the post-study survey, represented in the following heatmap. All participants were compensated for their time at £7.50 per hour, with a mean time of about 16 minutes. Figure B.2 shows a heat map of the demographics for these 84 participants.

### B.3.3 Response time

As a secondary explanatory variable response time is considered. Response time is first log transformed to remove its right skew. The same modeling procedure is repeated for

**Table B.3:** Model performance regressing on log response time [seconds],  $\widehat{Y}_2$  random effect models. Conditional  $R^2$  includes the random effects, while marginal does not. The model  $\alpha \times \beta + \gamma + \delta$  model is selected to examine further as it has relatively high marginal  $R^2$  while having much less complexity than the complete interaction model.

| Fixed effects | No. levels | No. terms | AIC         | BIC         | R2 cond.   | R2 marg.     | RMSE         |
|---------------|------------|-----------|-------------|-------------|------------|--------------|--------------|
| a             | 1          | 3         | <b>1448</b> | <b>1475</b> | 0.645      | 0.007        | 0.553        |
| a+b+c+d       | 4          | 8         | 1467        | 1516        | 0.647      | 0.017        | 0.552        |
| a*b+c+d       | 5          | 12        | 1474        | 1541        | 0.656      | 0.024        | 0.548        |
| a*b*c+d       | 8          | 28        | 1488        | 1627        | 0.673      | 0.054        | 0.536        |
| a*b*c*d       | 15         | 54        | 1537        | 1792        | <b>0.7</b> | <b>0.062</b> | <b>0.523</b> |

**Table B.4:** Model coefficients for log response time [seconds]  $\widehat{Y}_2 = \alpha \times \beta + \gamma + \delta$ , with factor = pca, location = 0/100%, shape = EEE, and dim = 4 held as baselines. Location = 50/50% is the fixed term with the strongest evidence and takes less time. In contrast, the interaction term location = 50/50%:shape = EEV has the most evidence and takes much longer on average.

|                             | Estimate | Std. Error | df    | t value | Pr(> t ) |     |
|-----------------------------|----------|------------|-------|---------|----------|-----|
| (Intercept)                 | 2.71     | 0.14       | 42.6  | 19.06   | 0.000    | *** |
| <b>Visual</b>               |          |            |       |         |          |     |
| Visualgrand                 | -0.23    | 0.12       | 567.6 | -1.97   | 0.049    | *   |
| Visualradial                | 0.16     | 0.12       | 573.5 | 1.34    | 0.181    |     |
| <b>Fixed effects</b>        |          |            |       |         |          |     |
| Location33/66%              | 0.05     | 0.14       | 40.9  | 0.34    | 0.737    |     |
| Location50/50%              | -0.05    | 0.14       | 42.1  | -0.35   | 0.729    |     |
| ShapeEEV                    | -0.15    | 0.09       | 8.3   | -1.61   | 0.145    |     |
| Shapebanana                 | -0.13    | 0.09       | 8.3   | -1.42   | 0.192    |     |
| Dim6                        | 0.14     | 0.08       | 8.3   | 1.90    | 0.093    |     |
| <b>Interactions</b>         |          |            |       |         |          |     |
| Visualgrand:Location33/66%  | 0.24     | 0.18       | 580.9 | 1.34    | 0.181    |     |
| Visualradial:Location33/66% | -0.24    | 0.18       | 582.4 | -1.32   | 0.188    |     |
| Visualgrand:Location50/50%  | 0.12     | 0.18       | 578.6 | 0.69    | 0.491    |     |
| Visualradial:Location50/50% | 0.05     | 0.18       | 584.4 | 0.25    | 0.800    |     |

this response. 1) Compare the performance of a battery of all additive and multiplicative models. Table B.3 shows the higher level performance of these models over increasing model complexity. 2) Select the model with the same effect terms,  $\alpha \times \beta + \gamma + \delta$ , with relatively high conditional  $R^2$  without becoming overly complex from interaction. The coefficients of this model are displayed in Table B.4.

## B.4 Chapter 5, cheem links

## APPENDIX B. SUPPLEMENTARY MATERIAL

---

| Content   | Link  |
|---|---|
| cheem vignette, Getting started with cheem                                  | <a href="https://shorturl.at/kqsF0">https://shorturl.at/kqsF0</a>                   |
| cheem documentation website   | <a href="https://nspyri.on.github.io/cheem/">https://nspyri.on.github.io/cheem/</a> |
| Penguins cheem radial tour, Figure<br><a href="#">5.4</a>                   | <a href="https://vimeo.com/666431172">https://vimeo.com/666431172</a>               |
| Chocolates, cheem radial tour, Figure<br><a href="#">5.6</a>                | <a href="https://vimeo.com/666431143">https://vimeo.com/666431143</a>               |
| Chocolates2, cheem radial tour, Figure<br><a href="#">5.7</a>               | <a href="https://vimeo.com/666431148">https://vimeo.com/666431148</a>               |
| FIFA 2020, cheem radial tour, Figure<br><a href="#">5.8</a>                 | <a href="https://vimeo.com/666431163">https://vimeo.com/666431163</a>               |
| North Ames housing prices, cheem radial tour, Figure<br><a href="#">5.9</a> | <a href="https://vimeo.com/666431134">https://vimeo.com/666431134</a>               |