blinded_abstract

Introduction

Missing data is ubiquitous in data analysis. ggplot2, an implementation of the grammar of graphics is an incredibly popular way to produce data visualisations does not currently support missing data. Principles of tidy data states that each row is an observation and each column is a variable, which makes it easy and consistent to perform data manipulation and wrangling. However, there are currently no guidelines for representing additional missing data structures in a tidy format. This paper describes approaches for exploring missing data structure with minimal deviation from the common workflows of ggplot and tidy data structures.