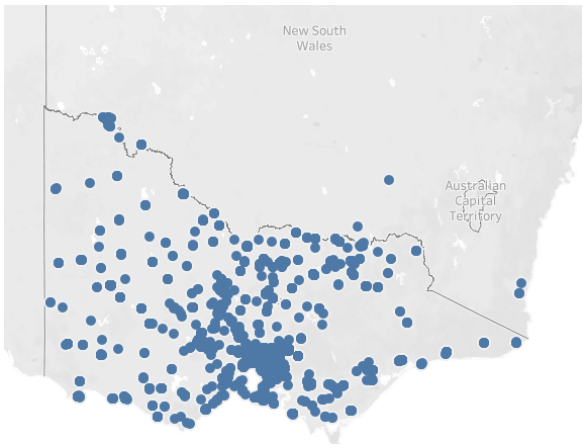
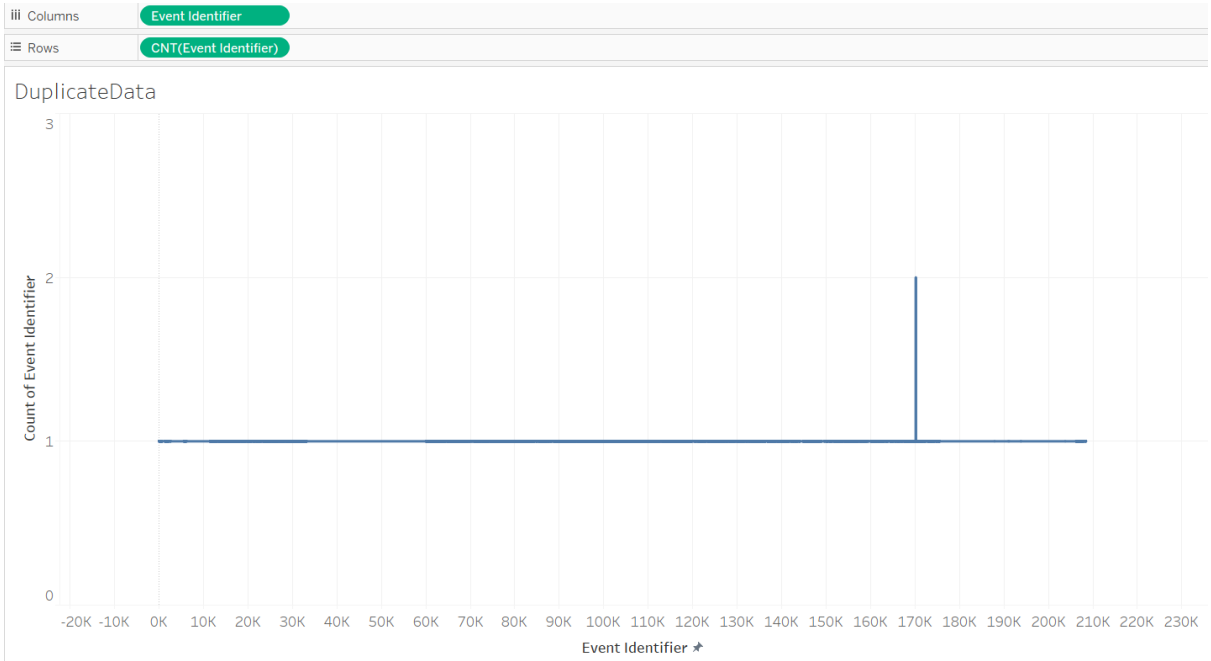
## 1. Data loading, checking and cleaning

The data is supposed to be of events in the state of Victoria, but some of the data is not in Victoria, as seen when plotting the latitude and longitude coordinates of the data. Remove them as they are irrelevant to us.




Not shown is some data appearing in the UK.

Some events are duplicated, as seen when plotting the number of event identifiers against the event identifiers. Remove duplicates, as they are clearly errors.

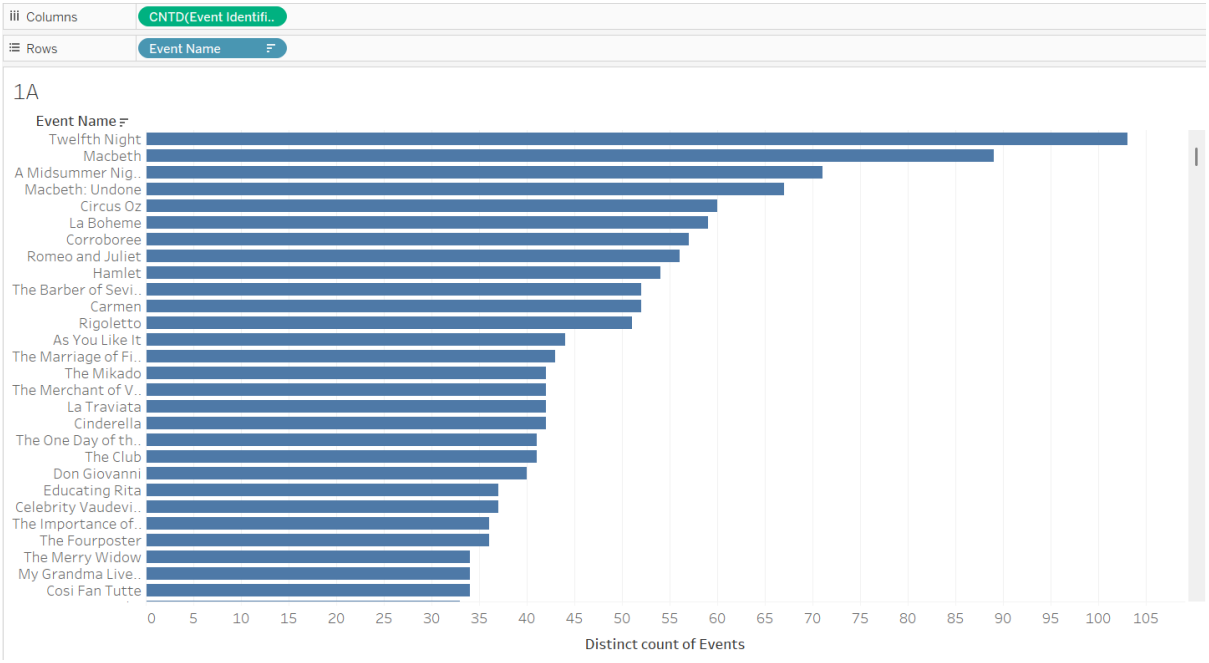



“First date” was in a string format and had to be converted to date format.

The above screenshots were taken before data cleaning. The rest of the screenshots in the report below were taken after data cleaning. All the visualizations in this report are based on a dataset provided by AusStage. (n.d.). Retrieved February 27, 2025, from http://www.ausstage.edu.au

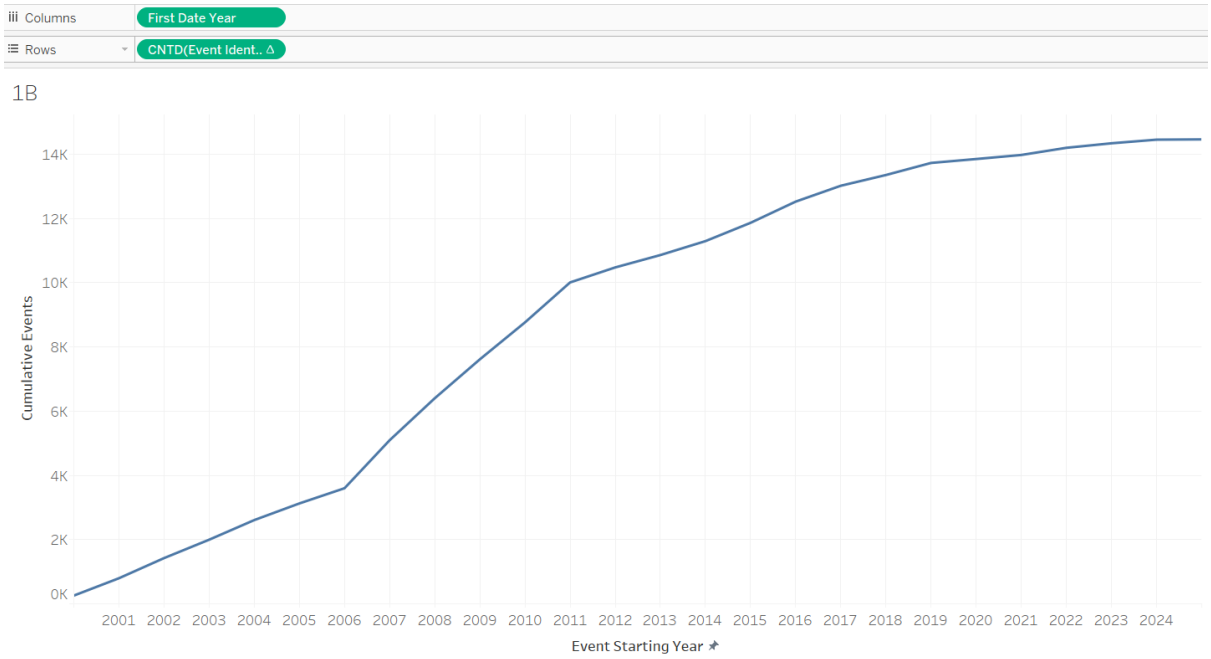
## 2. Data Exploration and Presentation

### A. What are the most common event names?



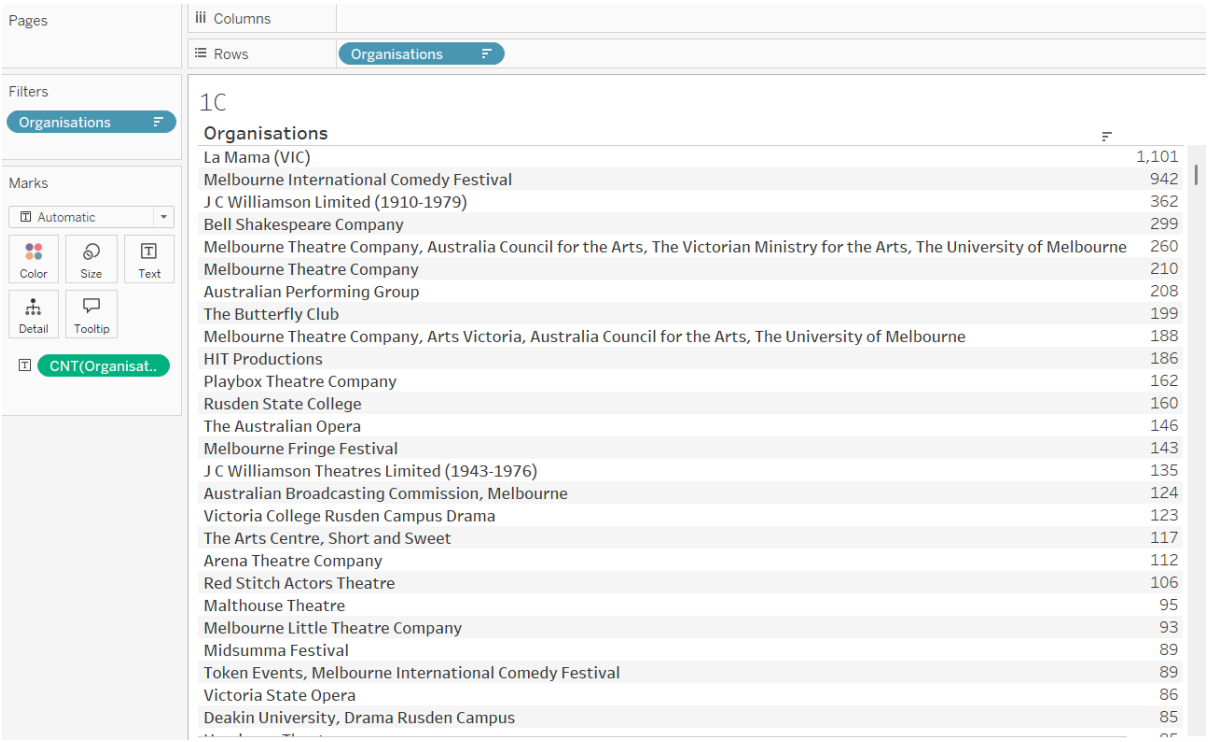

The most common event names can be visualized by plotting the event names against the number of events, identified by their event ID.

### B. How many events started each year over the last 25 years?




The number of events started each year can be identified by plotting the cumulative number of events, identified by their event ID, against the year they first started. This is then filtered to only include data since the year 2000.

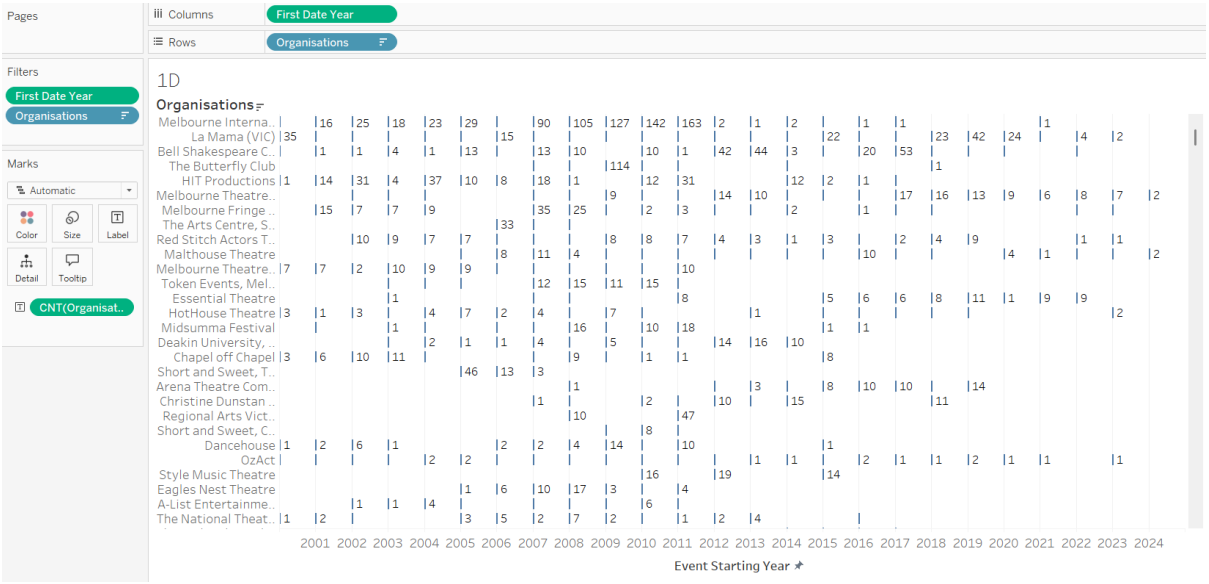
## C. How many events were run or performed by each organisation?




| Organisations | |
|---|---|
| La Mama (VIC) | 1,101 |
| Melbourne International Comedy Festival | 942 |
| J C Williamson Limited (1910-1979) | 362 |
| Bell Shakespeare Company | 299 |
| Melbourne Theatre Company, Australia Council for the Arts, The Victorian Ministry for the Arts, The University of Melbourne | 260 |
| Melbourne Theatre Company | 210 |
| Australian Performing Group | 208 |
| The Butterfly Club | 199 |
| Melbourne Theatre Company, Arts Victoria, Australia Council for the Arts, The University of Melbourne | 188 |
| HIT Productions | 186 |
| Playbox Theatre Company | 162 |
| Rusden State College | 160 |
| The Australian Opera | 146 |
| Melbourne Fringe Festival | 143 |
| J C Williamson Theatres Limited (1943-1976) | 135 |
| Australian Broadcasting Commission, Melbourne | 124 |
| Victoria College Rusden Campus Drama | 123 |
| The Arts Centre, Short and Sweet | 117 |
| Arena Theatre Company | 112 |
| Red Stitch Actors Theatre | 106 |
| Malthouse Theatre | 95 |
| Melbourne Little Theatre Company | 93 |
| Midsumma Festival | 89 |
| Token Events, Melbourne International Comedy Festival | 89 |
| Victoria State Opera | 86 |
| Deakin University, Drama Rusden Campus | 85 |

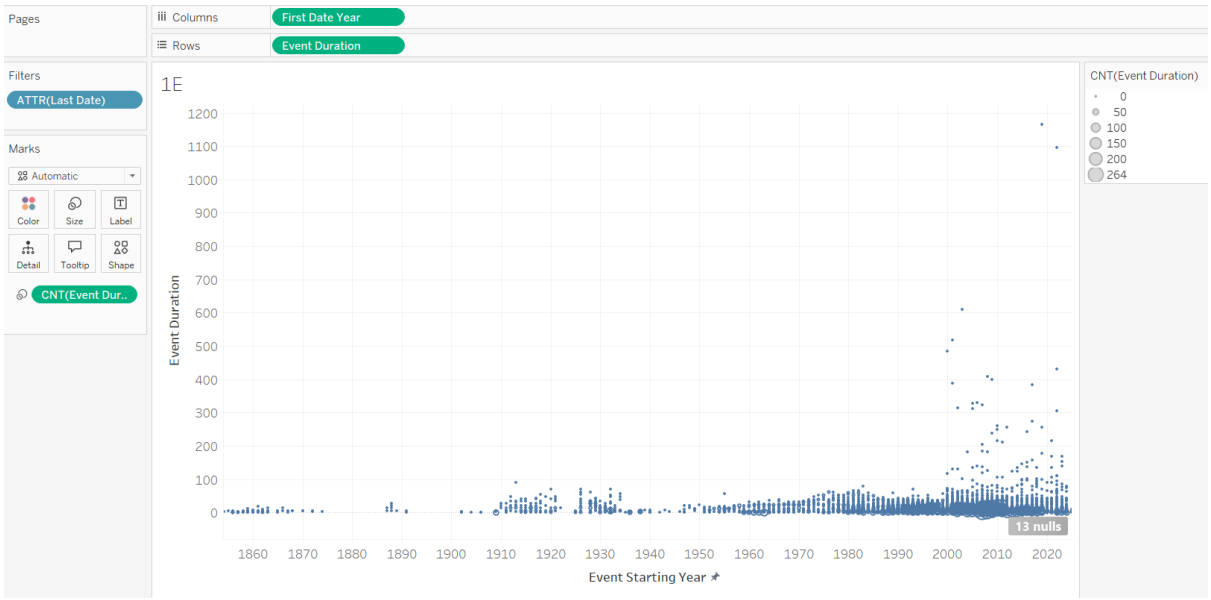
The number of events per organisation can be measured by plotting the number of times each organisation shows up in the data. This data is filtered to remove null values and then ordered to show the organisations with the most events run first.

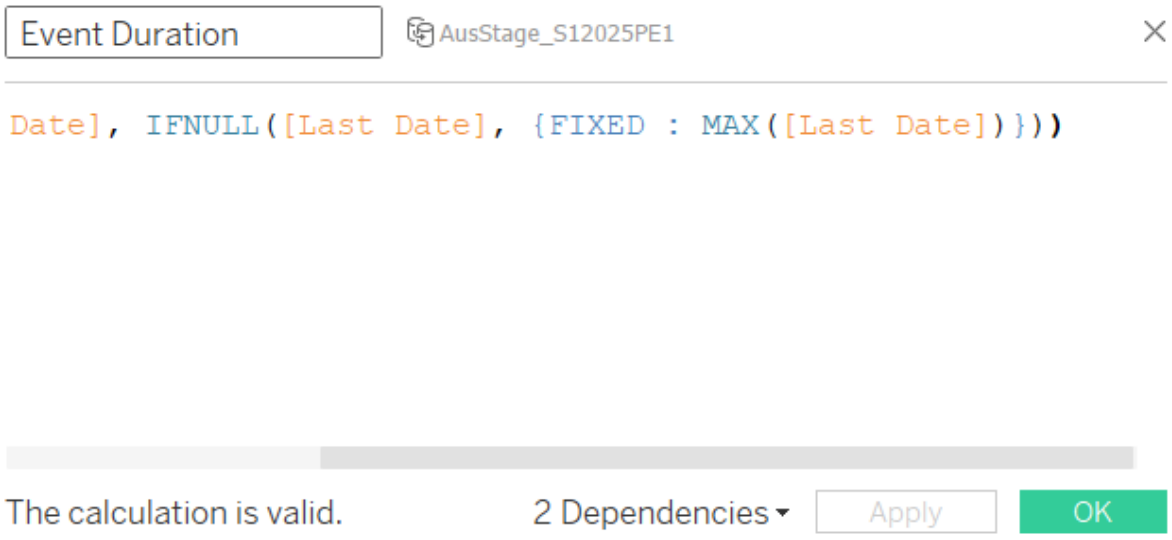
## D. How many organisations started events each year over the last 25 years?




As in C, the data is then plotted against the year the events started and filtered to only show data from the year 2000 onwards to show the number of events per organisation per year. The data is sorted according to most events since the year 2000 first.

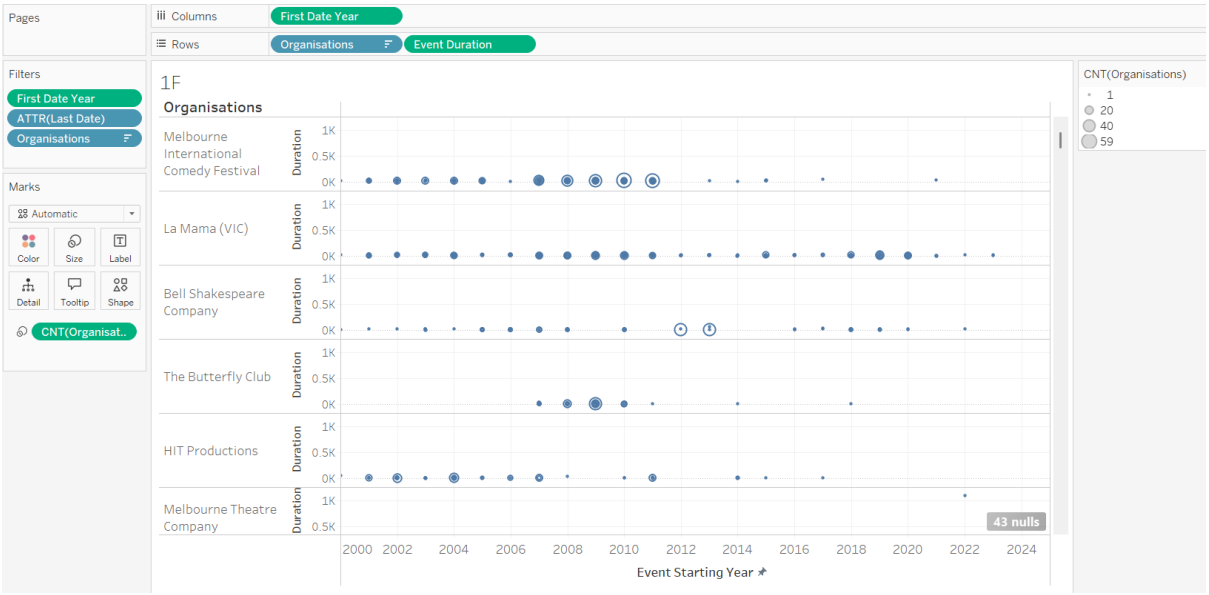
## E. How long did each event run for?




The event duration is computed by subtracting the event start date from the event end date. The event duration is then plotted against the event starting year, and filtered to remove null end dates, to prevent them from skewing the data. Clusters of high event density are shown as larger for visual clarity.

F. Based on the visualisations and findings for A-E, is it possible for you to now explain who (i.e., Organisation) ran or performed what events over the last 25 years?




Yes, question D shows which organisations started the events over the last 25 years. To also show how long the events organized by each organisation each year were, the duration can be added as an additional parameter as shown above.

## References

AusStage. (n.d.). AusStage. Retrieved February 27, 2025, from http://www.ausstage.edu.au