Shaping the Data

Lars Verspohl

VISUALIZATION DESIGNER AND DEVELOPER

@lars_vers   www.datamake.io
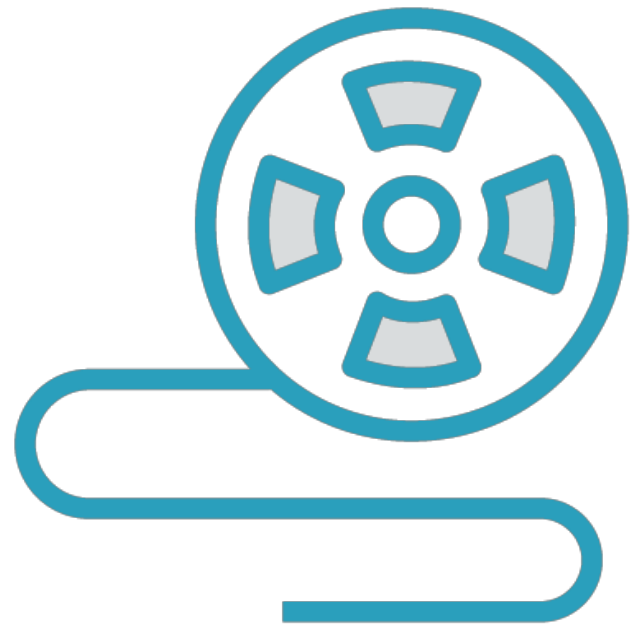
Overview

Data overview

Loading the data

Cleaning, shaping, filtering

Project structure

Data Overview

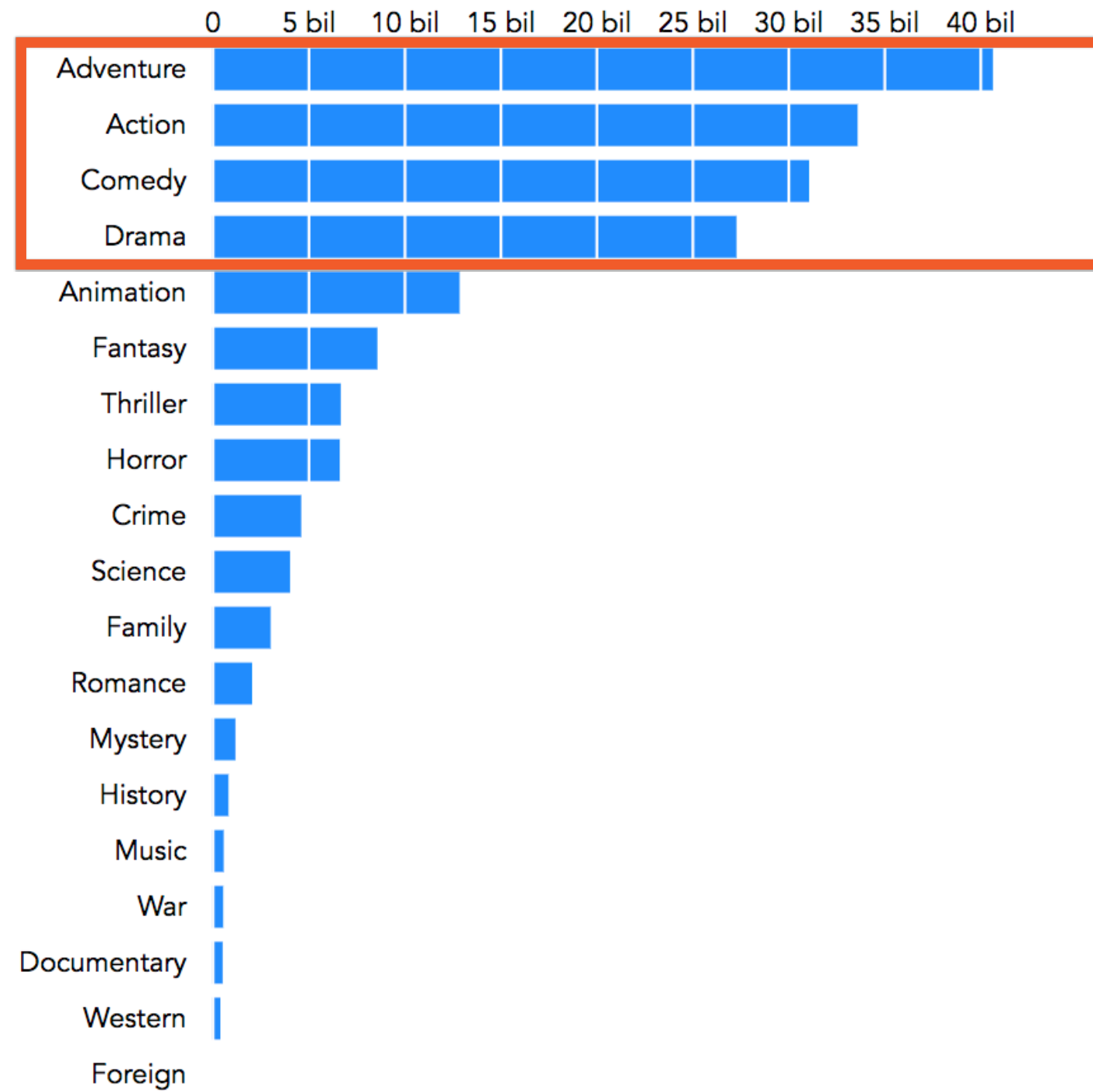**Movies**

2000
-
2009

Defined period

Lord
of the
Rings

Finding
Nemo

Avatar

Harry
Potter

Blockbuster

Financial success

Total revenue by genre in $US

Films w/ budget and revenue figures, 2000-2009

Genre	Revenue
	0 5 bil 10 bil 15 bil 20 bil 25 bil 30 bil 35 bil 40 bil
Adventure	████████████████████████████████████ (~41 bil)
Action	████████████████████████████████ (~34 bil)
Comedy	██████████████████████████████ (~31 bil)
Drama	█████████████████████████ (~26 bil)
Animation	████████████ (~13 bil)
Fantasy	█████████ (~9 bil)
Thriller	████████ (~8 bil)
Horror	████████ (~8 bil)
Crime	██████ (~6 bil)
Science	█████ (~5 bil)
Family	████ (~4 bil)
Romance	██ (~2 bil)
Mystery	█ (~1.5 bil)
History	█ (~1 bil)
Music		(~0.5 bil)
War		(~0.5 bil)
Documentary		(~0.5 bil)
Western		(~0.5 bil)
Foreign	

Total revenue by genre in $US

Films w/ budget and revenue figures, 2000-2009

	0	5 bil	10 bil	15 bil	20 bil	25 bil	30 bil	35 bil	40 bil

Adventure

Action

Comedy

Drama

Animation

Fantasy

Thriller

Horror

Crime

Science

Family

Romance

Mystery

History

Music

War

Documentary

Western

Foreign

Which film genres were most successful between 2000 and 2009?

Total revenue by genre in $US
Films w/ budget and revenue figures, 2000-2009
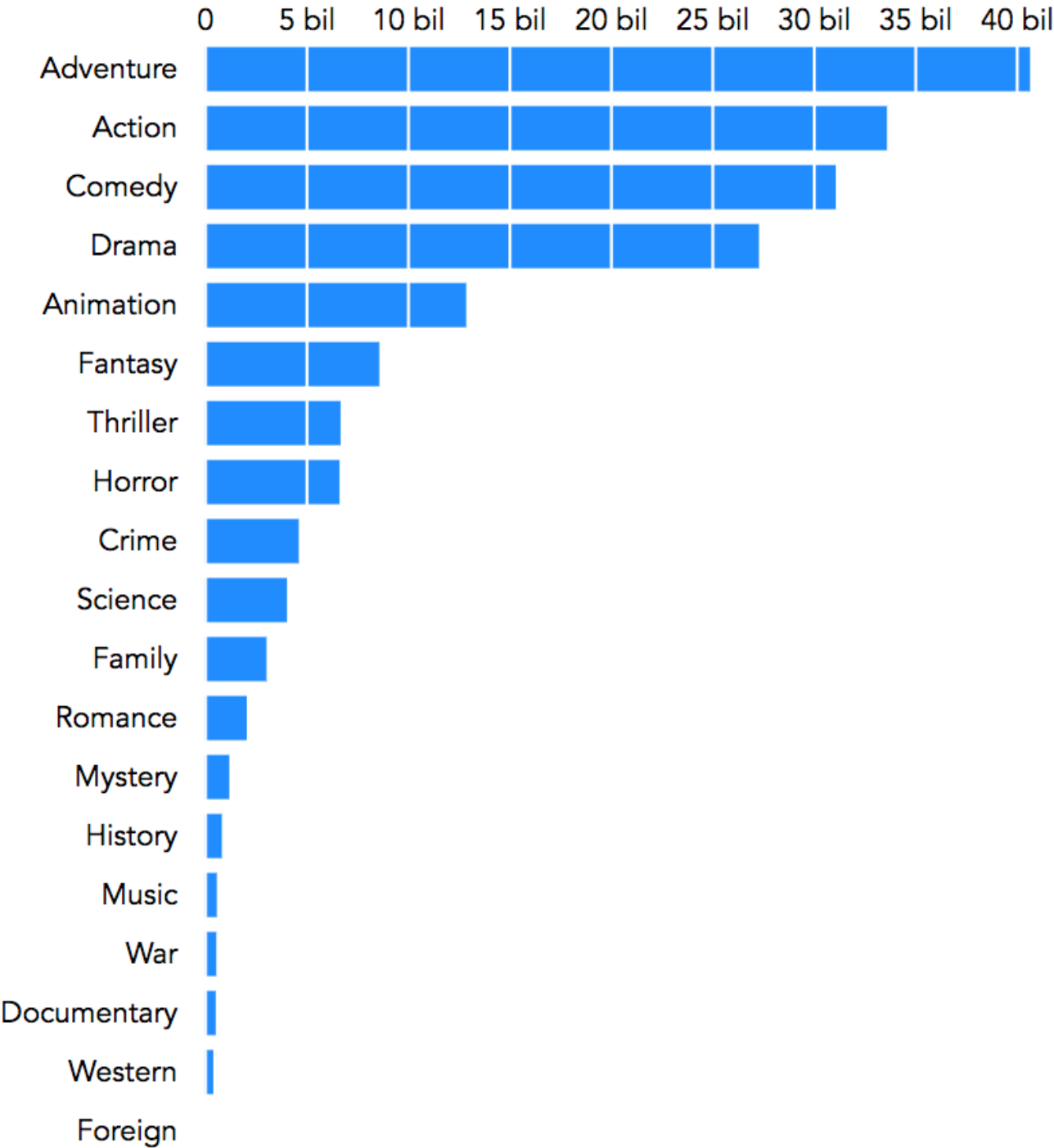
Genre	Revenue
Adventure	40
Action	33
Comedy	31
Drama	27
Animation	13
...	...

Total revenue by genre in $US

Films w/ budget and revenue figures, 2000-2009
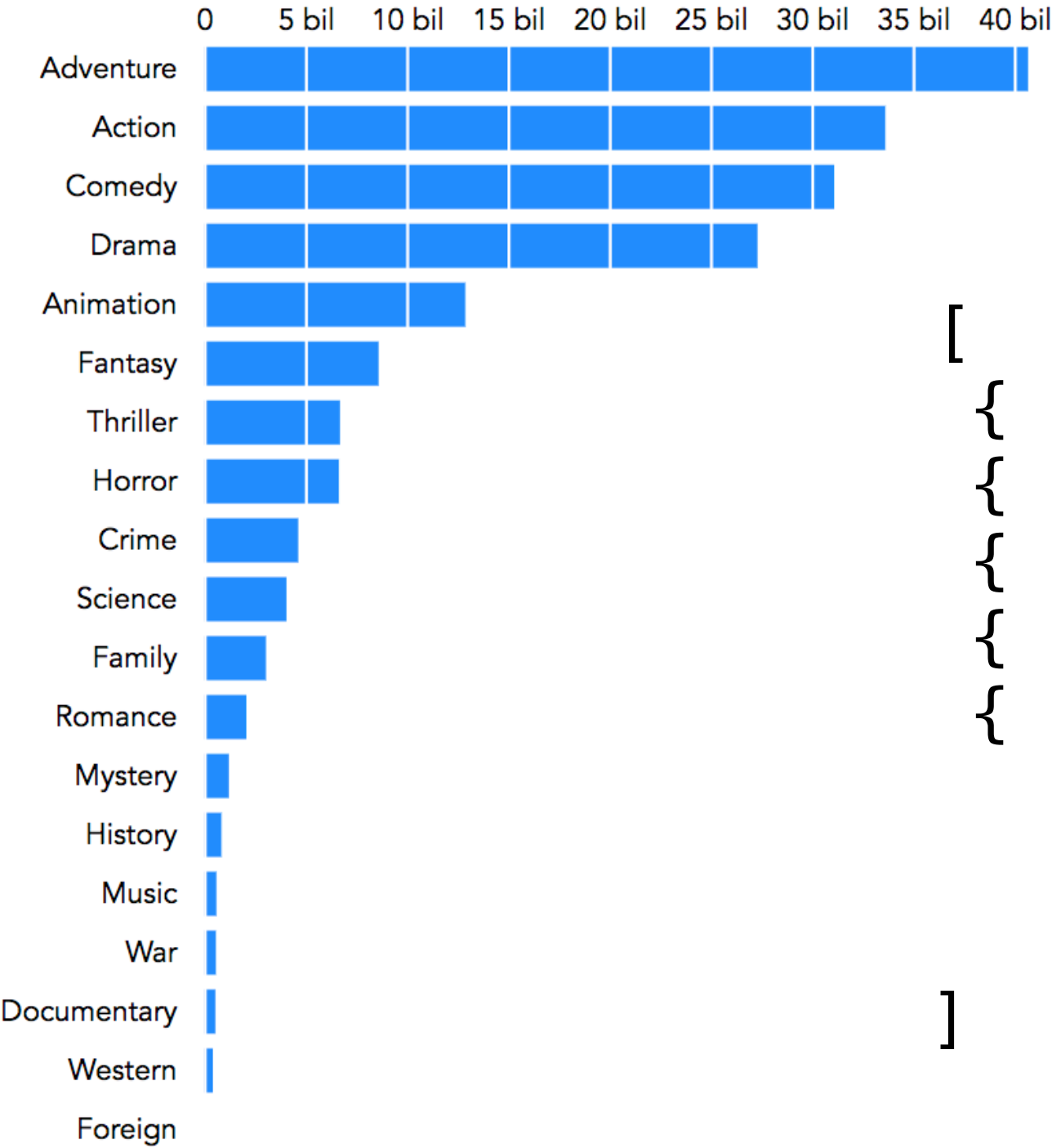

```
[
  { genre: "Adventure", revenue: 40 },
  { genre: "Action",    revenue: 33 },
  { genre: "Comedy",    revenue: 31 },
  { genre: "Drama",     revenue: 27 },
  { genre: "Animation", revenue: 13 },

  ...

]
```

d3.rollup(**data**, reducer, key)

Title	Genre	Year	...
Film 1	Adventure	2009	...
Film 2	Adventure	2002	...
Film 3	Comedy	2004	...
Film 4	Comedy	2003	...
Film 5	Comedy	2008	...
Film 6	Foreign	2000	...

`d3.rollup(data, reducer, `**`key`**`)`

Group data by:

Title	Genre	Year	...
Film 1	Adventure	2009	...
Film 2	Adventure	2002	...
Film 3	Comedy	2004	...
Film 4	Comedy	2003	...
Film 5	Comedy	2008	...
Film 6	Foreign	2000	...

`d3.rollup(data, `**`reducer`**`, key)`

Title	**Genre**	**Year**	**...**
Film 1	Adventure	2009	...
Film 2	Adventure	2002	...
Film 3	Comedy	2004	...
Film 4	Comedy	2003	...
Film 5	Comedy	2008	...
Film 6	Foreign	2000	...

Access to grouped subsets

Summary

Loading data with D3

- **Various formats**

- **Multiple sources**

Preparing data with D3 and JS

- **Type conversion**

- **New variables**

- **Filtering**

- **Transforming and aggregating**