

## LEGO SETS WITH THE MOST PIECES (AS OF 2021)



Figure 1: Lego sets with the most pieces

## Analysis

Analyze the above graph illustrating the Lego sets with the most pieces.

**Question: Is there one (or more) question addressed by the visualization?**

The question is clear and it is stated in the title of the visualization: what are the Lego sets with the highest number of pieces?

**Data:** Is the data quality appropriate?

Accuracy: data are absolute values representing the number of pieces, ranging from 4k to 12k, and they look reasonable.

Completeness: data are not complete, as only the top 15 Lego sets are considered.

Consistency: data are consistent, as they express an absolute value obtained by counting the number of pieces.

Currency: currency is appropriate, data are updated to the year 2021.

Credibility: the source is reported: it is a website (thecollector.io) and it seems trusted.

Understandability: the meaning of the values is quite clear (the number of pieces in each set).

Precision: precision up to the unit is appropriate for the task.

**Visual Proportionality: Are the values encoded in a uniformly proportional way?**

Not at all: the area of the bubbles seems not related to the corresponding value, but only to the “row” where they are located (smaller bubbles are in the bottom, bigger bubbles in the top). In any case, bubbles should be avoided because of Stevens’s law.

**Visual Utility: All the elements in the graph convey useful information?**

No, some elements can be removed: the Lego pieces in the background, the legend (in this case is useless), the bar below the title, the pictures of the Lego sets.

**Visual Clarity: Are the data in the graph clearly identifiable and understandable (properly described)?**

The colors associated with the Lego sets have no real meaning and they represent a clarity issue as they are a source of distraction for the reader. The order of the values is following the convention left-right and top-bottom, but this is not appropriate for a ranking (either use left-right or top-bottom).

**Design**

Design the visualization based on the following data structure.

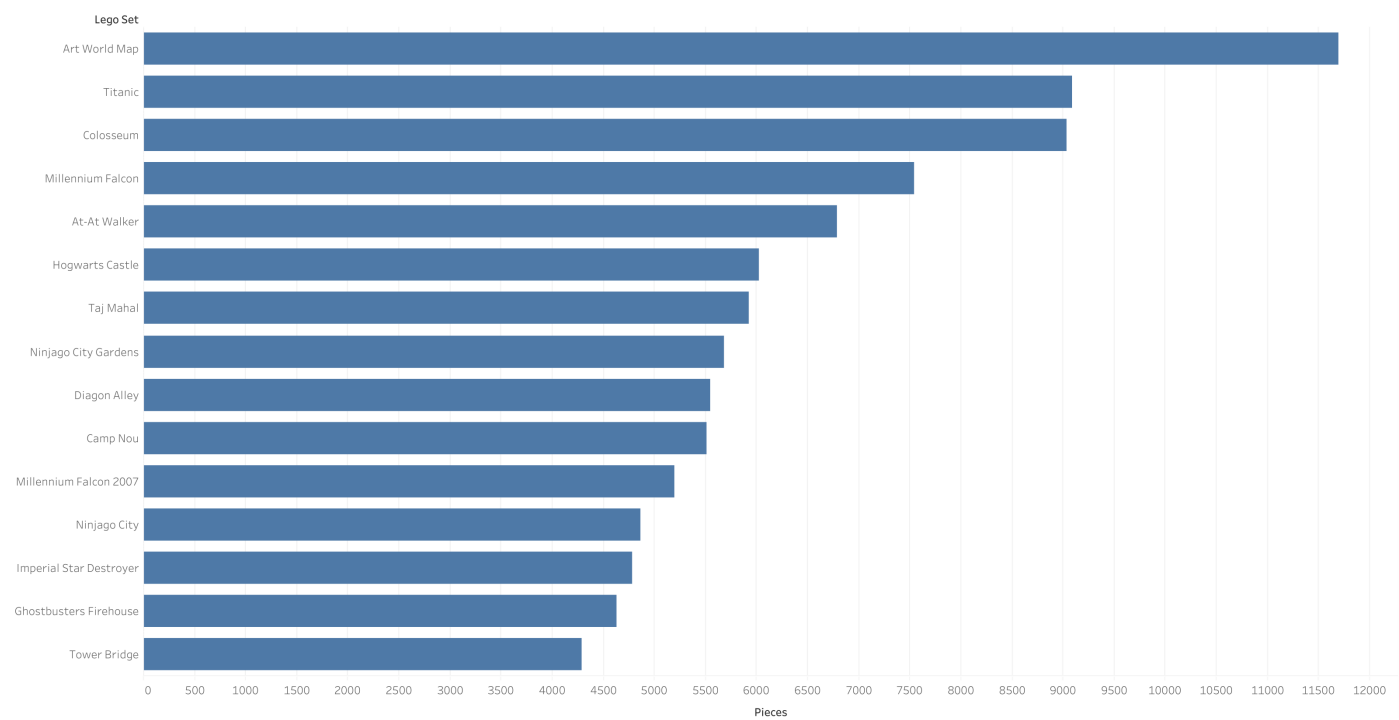
Field	Dim./Measure	Description
PIECES	Measure	The number of pieces in the Lego set
LEGO_SET	Dimension	The name of the Lego set
COLOR	Dimension	No real meaning

Design schema

Schema	Details
Columns:	SUM(PIECES)
Rows:	LEGO_SET
Graph type:	Bar
Color:	Default
Size:	Default
Label:	Default

Sketch of the resulting graph

Bar chart

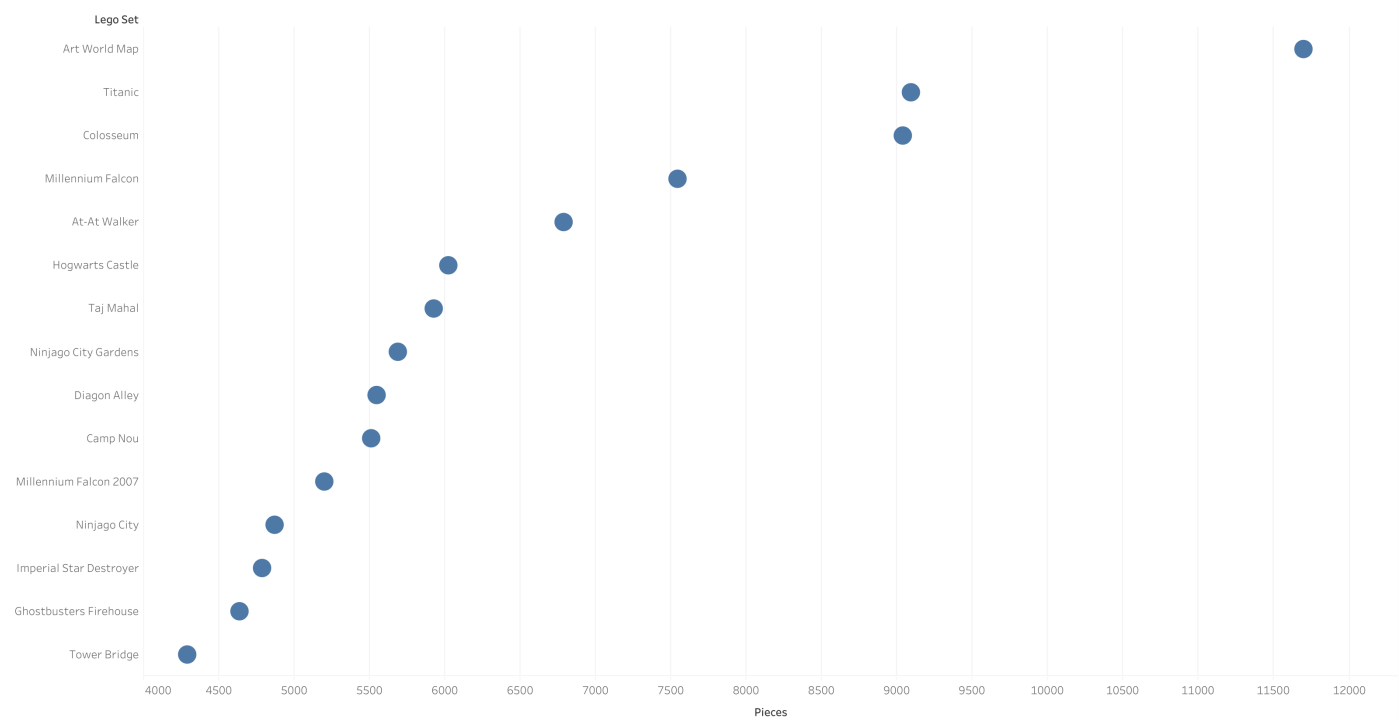


Design schema

Schema	Details
Columns:	SUM(PIECES)
Rows:	LEGO_SET
Graph type:	Circle
Color:	Default
Size:	Default
Label:	Default

Sketch of the resulting graph

Dot plot

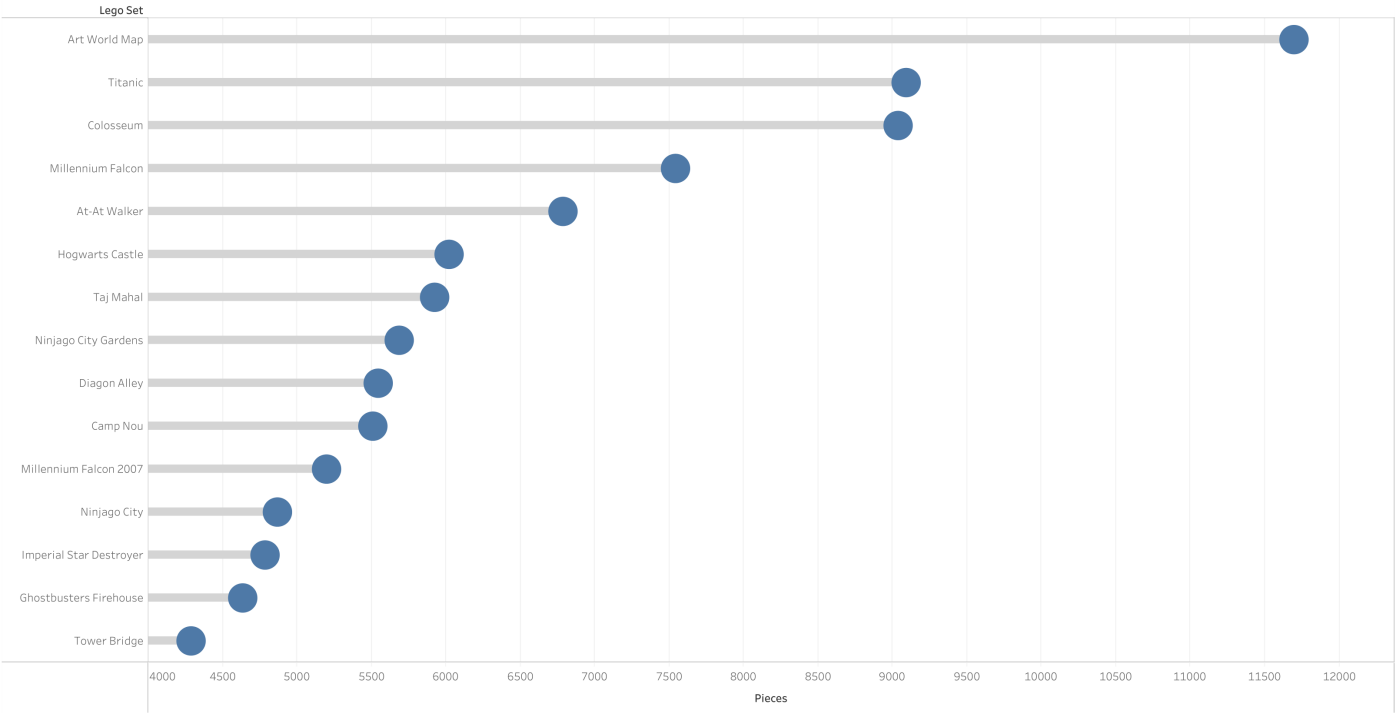


Design schema

Schema	Details
Columns:	SUM(PIECES), SUM(PIECES)
Rows:	LEGO_SET
Graph type:	Bar, Circle
Color:	Default
Size:	Default
Label:	Default

Sketch of the resulting graph

Lollypop



## Theory

Which of the following is a valid reason to use graphs instead of tables?

- To look up individual values
- To compare values over time
- To use very precise values
- To report more than one unit of measure
- *To reveal relationships among multiple values*