

# Data Visualization

Data visualization transforms data into stories your audience can understand instantly — it's easier to understand complex information when you can see it visualized.

A data-driven perspective adds depth to your research by revealing new angles and viewpoints. It improves analytical accuracy and uncovers hidden patterns and trends.

**But only if it's done right.**

The essential question: ***What does this visualization add to my research?***

Data visualization should tell compelling stories. It's not just about displaying data and showing beautiful pictures, it's about communicating insights that matter to your research objectives.

## **Fundamental Rules:**

### ***1. Answer 2 questions: Do I really need this visualization, and why?***

If you can clearly answer both questions, proceed.

### ***2. Match Visualization to Data Type***

The wrong chart type can mislead both you and your audience, even with accurate data.

### ***3. Prioritize Clarity Over Complexity***

Complex visualizations may confuse your audience—always aim for maximum clarity.

### ***4. Data is Part of Your Narrative***

Provide your audience with context, not just raw information.

# Data Visualization and Analysis Tools

## 1. [Rowboat](#)

One of the recently presented tools we're currently using in our projects for data analytics - Rowboat.

Rowboat is a web-based tool developed for quickly understanding large datasets, such as Excel spreadsheets and CSV files.

### ***Key features:***

1. Rowboat uses visualization to help users grasp the meaning of their data, inspiring new questions and insights.
2. Users can filter, search, explore, and navigate their data efficiently
3. Rowboat doesn't require installation; you only need to register.
4. Rowboat allows users to drag and drop files directly into the browser for instant analysis, but it keeps data private.
5. Provides immediate visual summaries of datasets without leaving the data interface.

Rowboat is designed to make data exploration accessible without requiring users to learn complex tools or write custom code.

Once you upload your dataset, you can see that statistics are immediately available, highlighting popular publishing houses, authors, languages, and more. This allows us to explore the data in-depth and makes it easy to build hypotheses and ask new questions.

Rowboat's standout feature is its dynamic filtering system combined with real-time statistical calculations, which creates a seamless data exploration experience. Once you've found that perfect slice of data you're looking for, you can easily export it and keep working with just that subset.

Additionally, Rowboat allows you to download visualizations in multiple formats, including PDF and PNG, for presentations or reports.

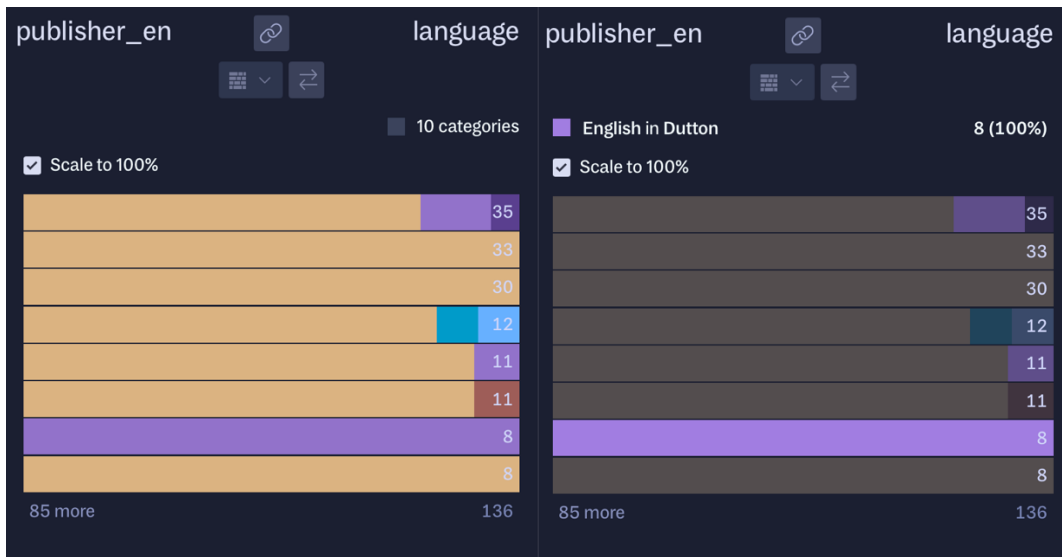
*Ru realm dataset, examples:*



*Countries of origin. 48% of published Tamizdat authors were born in the Soviet Union*



*Language of publications. Russian is the most used one; if translated, books were mostly published in English*



*Publishers and Translations. While some publishers only printed the original texts, others, like Dutton, only published translations*

Rowboat ru-realm-clean-xlsx.csv

author\_inifed Mikhail Bulgakov language English

y_of_b...	title_en	publisher_en	language	translator	year
	Filter by	Filter by	Filter by	Filter by	Filter by
	SORT BY A-Z COUNT	SORT BY A-Z COUNT	SORT BY A-Z COUNT	SORT BY A-Z COUNT	SORT BY A-Z COUNT
	The Master and Margarita 2	Grove 2	All entries are identical	Michael Glenny 3	1967 3
	Black Snow: A Theatrical Novel 1	Harcourt 1	Filtered to English	Mirra Ginsburg 2	1968 2
	Heart of a Dog 1	Harper & Row 1	Remove filter		
	The Heart of a Dog 1	Simon and Schuster 1	Hide this column		
	ct value 20%	4 distinct values 80%	1 distinct value 20%	2 distinct values 40%	2 distinct values 40%
181	The Master and Margarita	Grove	English	Mirra Ginsburg	1967
182	The Master and Margarita	Harper & Row	English	Michael Glenny	1967
183	Black Snow: A Theatrical Novel	Simon and Schuster	English	Michael Glenny	1967
184	Heart of a Dog	Grove	English	Mirra Ginsburg	1968
185	The Heart of a Dog	Harcourt	English	Michael Glenny	1968

*Filtered and grouped data to see only Bulgakov's publications in English. Can see the same people working on translations, which raises the question of the translator as an actor in the publication process*

Rowboat ru-realm-clean-xlsx.csv language French

author_inifed	title_en	publisher_en	language	translator	year
Filter by	Filter by	Filter by		Filter by	Filter by
SORT BY A-Z COUNT	SORT BY A-Z COUNT	SORT BY A-Z COUNT		SORT BY A-Z COUNT	SORT BY A-Z COUNT
Varlam Shalamov 2	Article 58: mémoires du prisonnier Chalamov 1	Denoël 1		Jean-Jacques Marie 1	1969
Joseph Brodsky 1	Collines et autres poèmes 1	Gallimard 1		Olivier Simon (Jean-Jacques Marie); K... 1	1966
	Récits de Kolyma 1	Éditions du Seuil 1			
			All entries are identical		
			Filtered to French		
			Remove filter		
			Hide this column		
2 distinct values 67%	3 distinct values 100%	3 distinct values 100%	1 distinct value 33%	1 empty entry 33% 2 distinct values 100%	2 distinct values 100%
208 Joseph Brodsky	Collines et autres poèmes	Éditions du Seuil	French	Jean-Jacques Marie	
273 Varlam Shalamov	Article 58: mémoires du prisonnier Chalamov	Gallimard	French		
276 Varlam Shalamov	Récits de Kolyma	Denoël	French	Olivier Simon (Jean-Jacques Marie); Katia Ke	

Publications in French. The same translator, but under a nickname for Shalamov's text, uses the real name for Brodsky's translation

Rowboat ru-realm-clean-xlsx.csv translator Max Hayward

author_inifed	title_en	translator	publisher_en	year
Filter by	Filter by		Filter by	Filter by
SORT BY A-Z COUNT	SORT BY A-Z COUNT		SORT BY A-Z COUNT	SORT BY A-Z COUNT
Nadezhda Mandelstam 2	Hope Abandoned 1		Atheneum 2	1960 1
Aleksandr Gladkov 1	Hope Against Hope 1		Collins & Harvill 1	1966 1
Andrei Sinyavsky 1	Meetings with Pasternak. A Memoir 1		Harper & Row 1	1970 1
	On Trial: The Soviet State Versus Abram Tertz ... 1		Pantheon Books 1	1974 1
	The Trial Begins 1			1977 1
		All entries are identical		
		Filtered to Max Hayward		
		Remove filter		
		Hide this column		
1 empty entry 20% 3 distinct values 75%	5 distinct values 100%	1 distinct value 20%	4 distinct values 80%	5 distinct values 100%
187 Andrei Sinyavsky	On Trial: The Soviet State Versus Abram Tertz and Nik	Max Hayward	Harper & Row	1966
199 Nadezhda Mandelstam	The Trial Begins	Max Hayward	Pantheon Books	1960
220 Nadezhda Mandelstam	Hope Against Hope	Max Hayward	Atheneum	1970
222 Nadezhda Mandelstam	Hope Abandoned	Max Hayward	Atheneum	1974
241 Aleksandr Gladkov	Meetings with Pasternak. A Memoir	Max Hayward	Collins & Harvill	1977

Texts translated by Max Hayward for different publishing houses over the years. Allows us to explore his relationships both with authors and publishers

## 2. Cosmograph. Network visualization

Network visualization (also called graph visualization) is the process of visually presenting networks of connected entities as links and nodes. Network analysis examines relationships between different entities, such as publishers and authors, authors and the languages they were translated into, and texts and their translators (as in the previous example, we saw several people translating the same piece).

The visual aspect is crucial because this type of visualization simplifies the complex nature of the network, making it easier to understand patterns that might be hard or impossible to notice in raw data.

### ***Why Cosmograph?***

Unlike other tools, Cosmograph is dynamic and interactive. When you select a node, it isolates it from other data entries, allowing you to view just that subset. This lets you focus on a specific cluster and analyze it in more detail without distractions.

Cosmograph has a clean, minimal interface. When you open it, you'll find data examples right away, which provide hints on how you should structure your data for this specific visualization tool. The simple and intuitive interface allows you to quickly manage and customize your visualization.

### ***How to use:***

1. Import your data in CSV format
2. Select the entity connections you want to analyze
3. Set one as the Source and another as the Target
4. Customize your visualization using the settings



### 3. Flourish. Maps.

Using a map visualization might help you communicate your data to your audience because it goes beyond just numbers and graphs — it provides a familiar reference point, making data more intuitive and easier to connect with by creating a narrative.

#### Visualizing Ru Realm

For Flow Map type visualization, we require two distinct datasets: the data itself and its associated metadata.

The first dataset includes key details such as:

- Author
- Author's country of birth
- Author's country of publication

The metadata dataset contains geolocation information, including:

- Location name
- Latitude
- Longitude

The metadata spreadsheet structure:

The first column lists locations — both authors' origin countries and countries of publication. Each location has its corresponding latitude and longitude coordinates, ensuring accurate mapping. This metadata structure is essential for creating geographic visualizations. The settings are very intuitive — you simply need to enter the corresponding column names: latitude to latitude, longitude to longitude.

Moving on to the data itself:

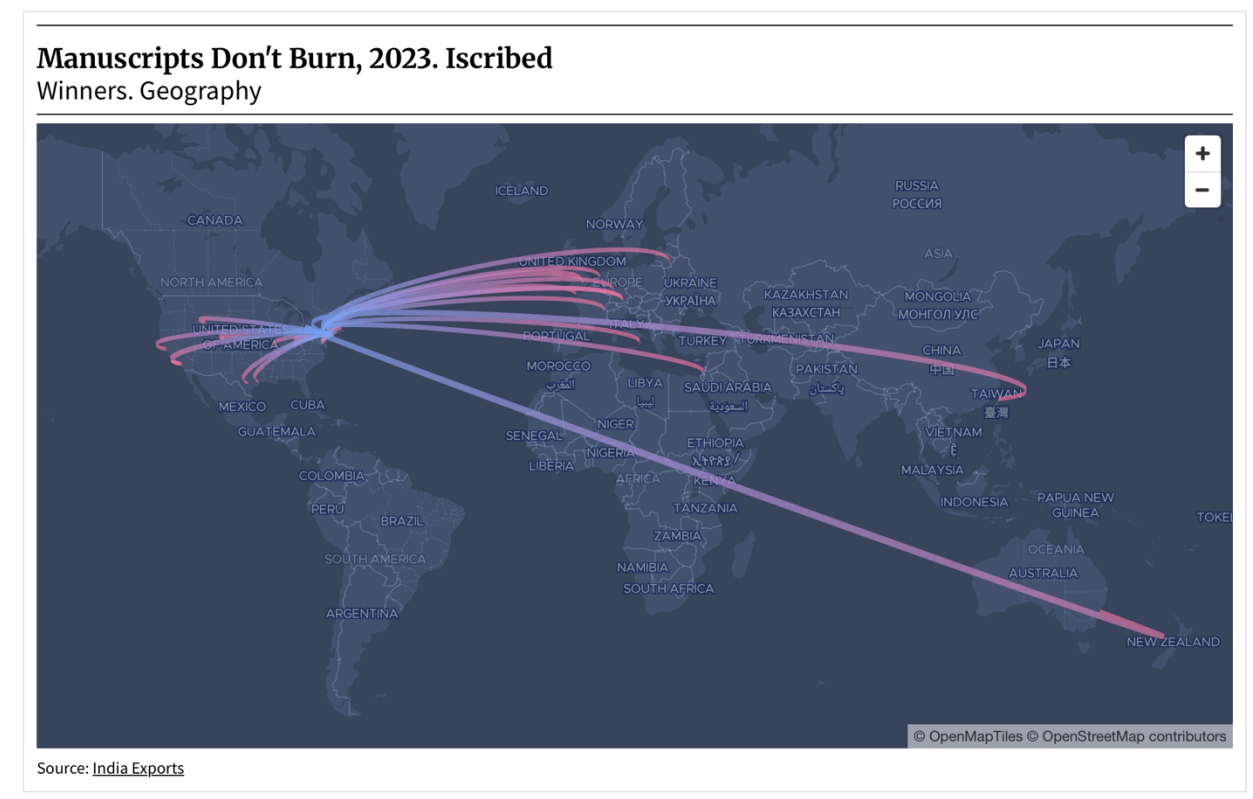
Since we are visualizing flows, we need to establish connections between two points: the origin and the destination. The corresponding fields in the settings are:

- Source location (the author's birth country)
- Destination location (the publication city)
- Filter (author)



Once mapped, we can see the connections between each author's origin country and their publication country. You can use the filter to explore the publication flow for individual authors, allowing you to trace how literary works move across geographic boundaries.

Flow visualization example (Tamizdat Auction 2023)



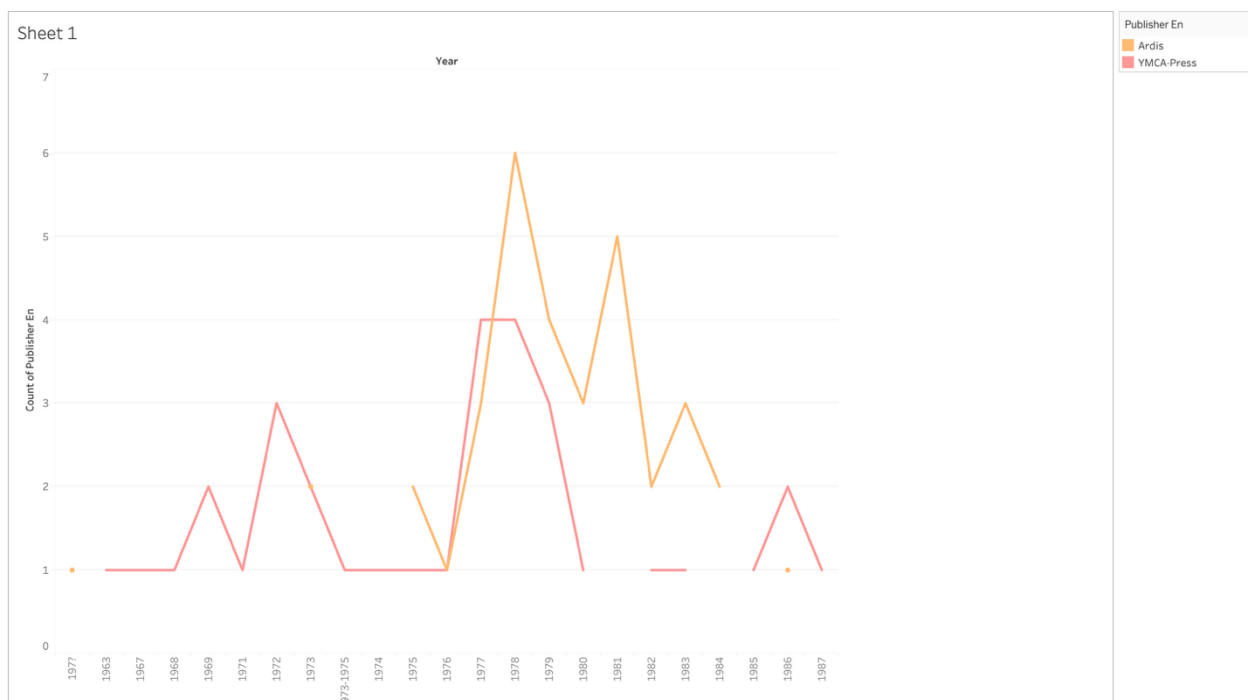
*Blue – Tamizdat Project location (source); Pink – Auction winners(destination)*

# Basic Chart Types

## 1. Line Charts

Connect distinct data points through straight lines to reveal trends, patterns, and changes. Perfect for showing how something changes over time.

Example:



Line chart. Number of published books by Ardis and YMCA-Press

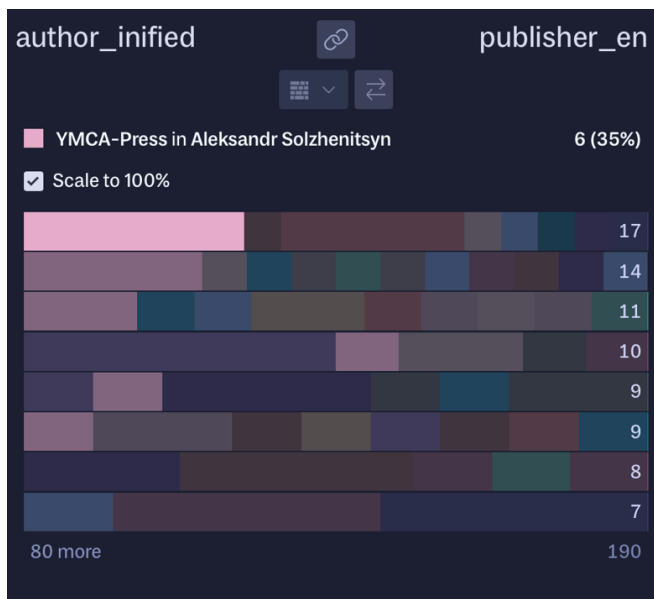
## 2. Bar Charts

Visually represent data using rectangular bars where the length of each bar corresponds proportionally to its value. There are various types, like horizontal bar chart, stacked bar chart, grouped bar chart.

Examples:



Horizontal bar chart. Authors by total count



Stacked bar chart. Authors published by the publishing house, % and total count

### 3. Pie Chart

A circular, statistical graphic that divides data into slices, where each slice represents a percentage or proportion of the whole.

Example:

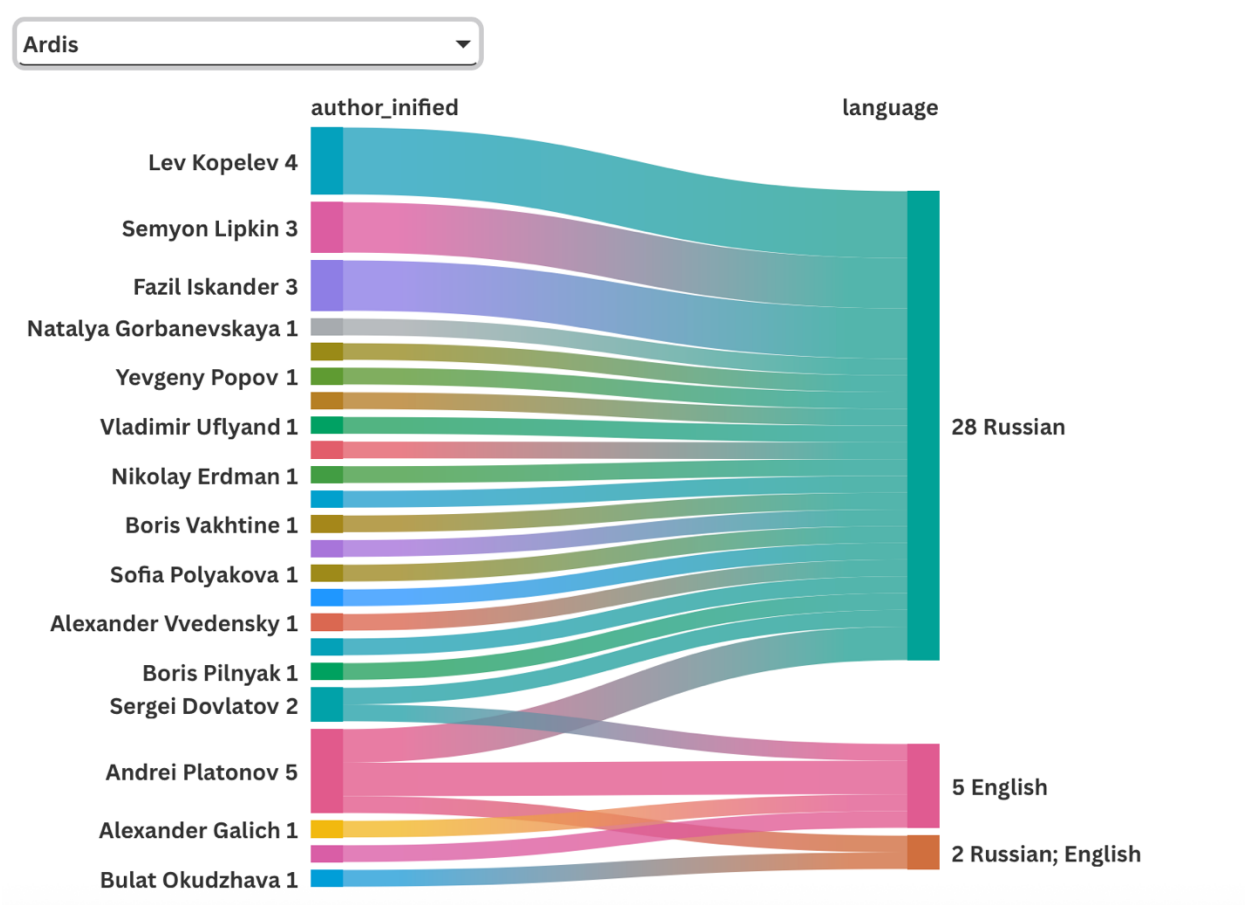


*Pie chart. Language by percentage*

#### 4. Sankey Diagrams

Flow charts that illustrate the movement of values from one set to another, with items connected using colored lines or arrows.

Example:



*Sankey Diagram. Ardis published authors and the languages their texts were published*