SEARCHING FOR MY NEXT CHART

MUHAMMAD NAKHAEE

PYDATA AMSTERDAM 2025

Work as an ai engineer @ Exact

Love photography and collecting cameras @merrygoroundofphotos

Exploring my creative side these days

Perfectionist!

This Talk

is **NOT** a tutorial on how to design a great chart

looks at “data visualization” as a creative form, not just a practical tool.

My “Creative” DataViz Evolution

My “Creative” DataViz Evolution

Why ggplot2

Intuitive (based on the grammar of graphics)

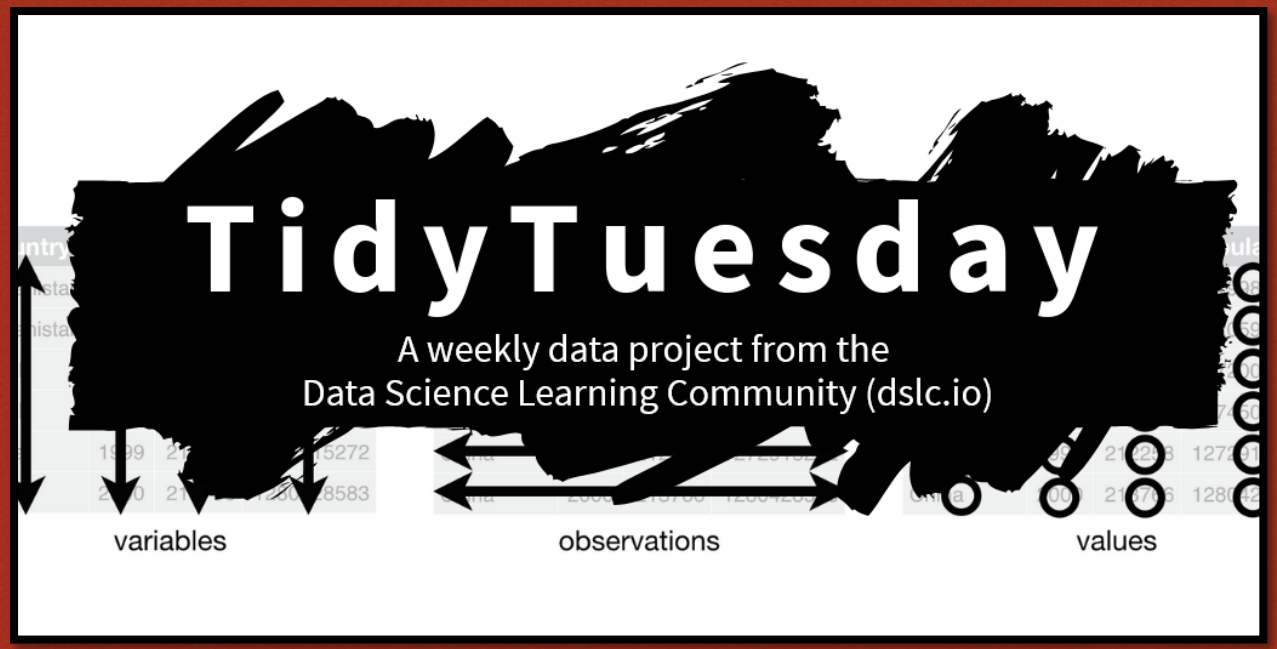
Powerful

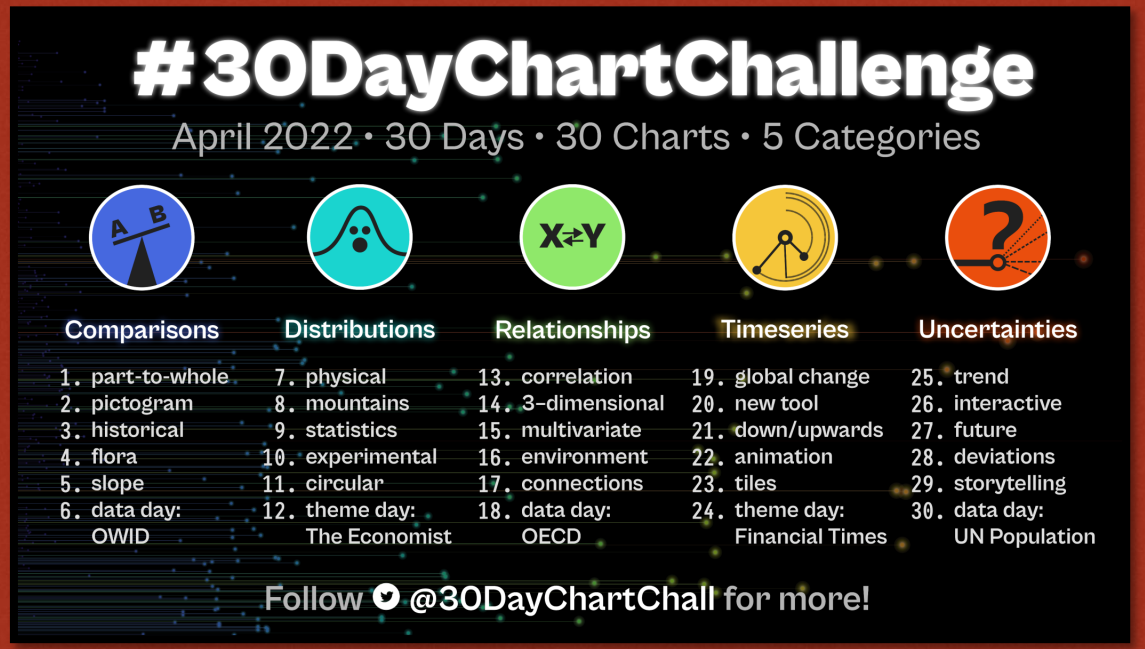
Flexible

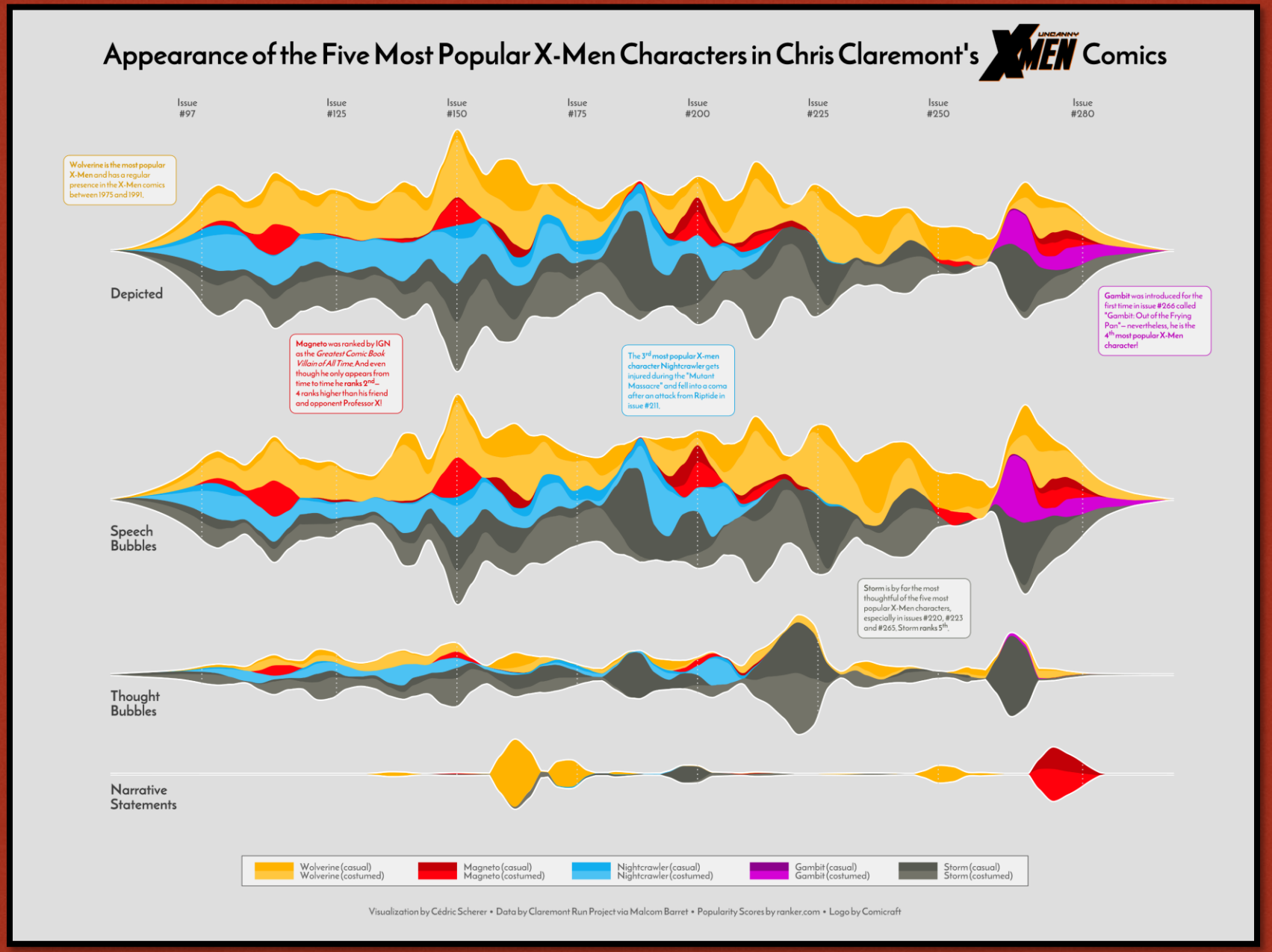
User base

Developer base

TidyTuesday



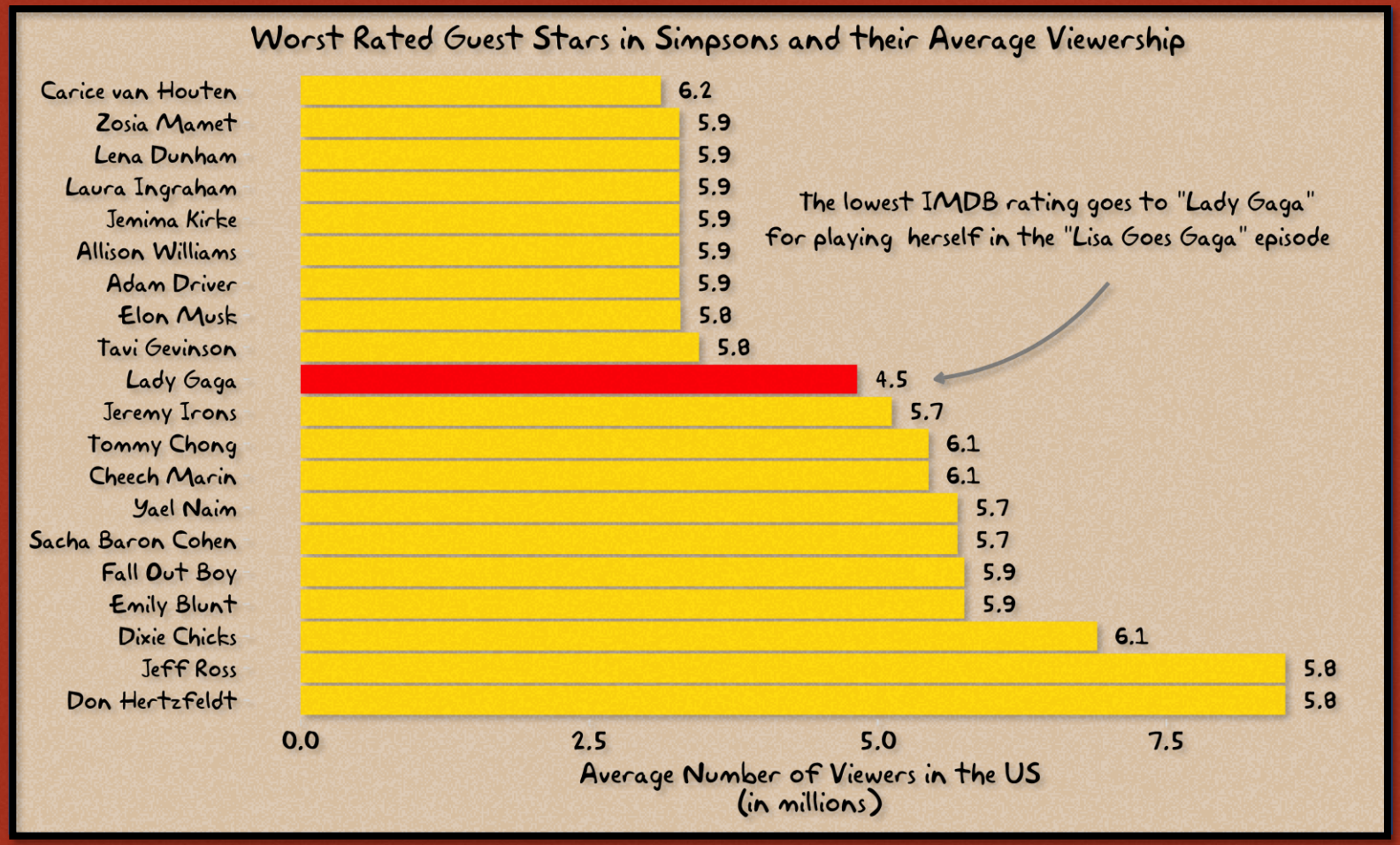




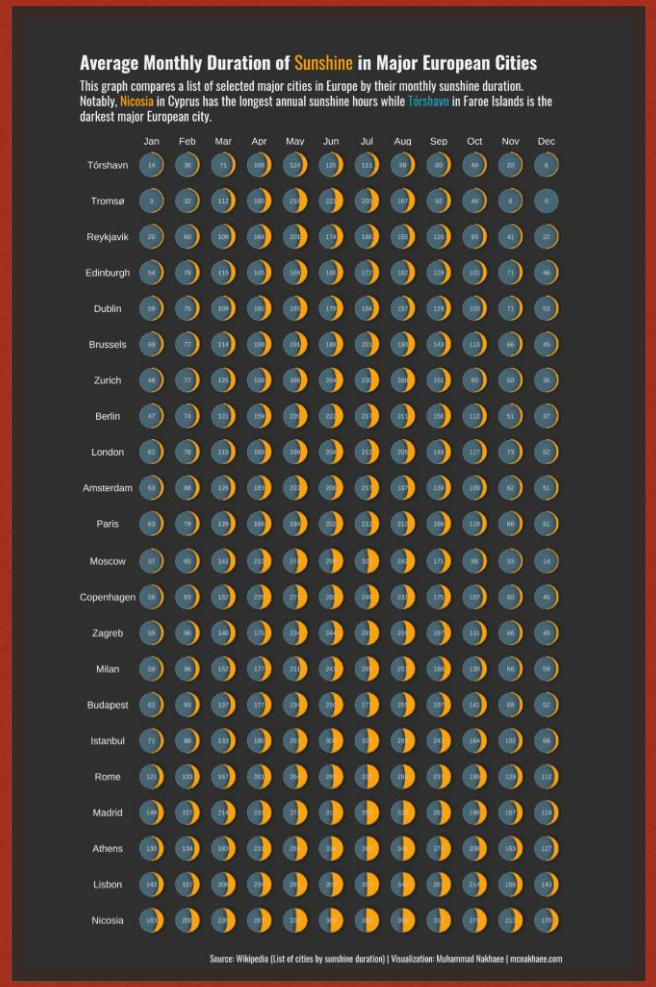
Cédric Scherer Uncanny X-Men by Claremont Run Project

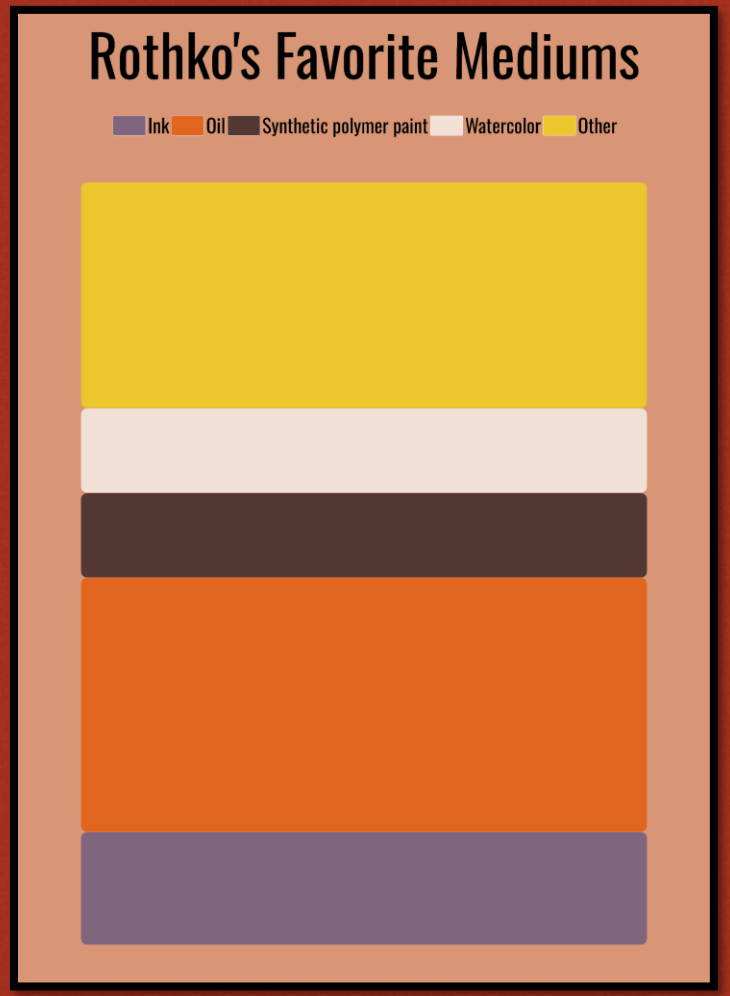
CEO departures, Georgios Karamanis

My First #TidyTuesday Submission



#30DayChartChallenge





Winning a Kaggle Prize

Data Analyst

Professor/Teacher

Research Scientist

Statistician

Data Scientist

ML Engineer

Data Engineer

Software Engineer

Manager

Main Missions:

1. Analyze and understand data to influence product or business decisions 2. Build and/or the data infrastructure that my business uses storing, analyzing, and operationalizing data

7,500-9,999 Median Compensation

Under 1 year Median Coding XP

Under 1 year Median ML XP

Python 9

SQL 4

R 5

Javascript 9

C++ 11

C 8

MATLAB 5

MySQL 7

Microsoft SQL Server 7

PostgreSQL 9

SQLite 9

Google Cloud BigQuery 5

Oracle Database 7

MongoDB 11

Scikit-learn 9

TensorFlow 10

Keras 12

Xgboost 9

PyTorch 11

LightGBM 11

Caret 3

Jupyter Notebook 7

Visual Studio Code (VSCode) 10

PyCharm 5

RStudio 4

Matplotlib 7

Seaborn 5

Plotly / Plotly Express 7

Ggplot / ggplot2 5

Tableau 2

Microsoft Power BI 2

Google Data Studio 3

Qlik Sense 2

Google Cloud Platform (GCP) 11

Amazon Web Services (AWS) 9

Microsoft Azure 9

Oracle Cloud 5

Linear or Logistic Regression 10

Decision Trees or Random Forests 10

Gradient Boosting Machines 9

Convolutional Neural Networks 11

Word embeddings vectors 12

Transformer language models 11

Encoder-decoder models 11

Contextualized embeddings 10

Image classification 12

Image segmentation methods 11

Object detection methods 12

General purpose image video tools 12

TensorFlow Extended (TFX) 9

MLflow 9

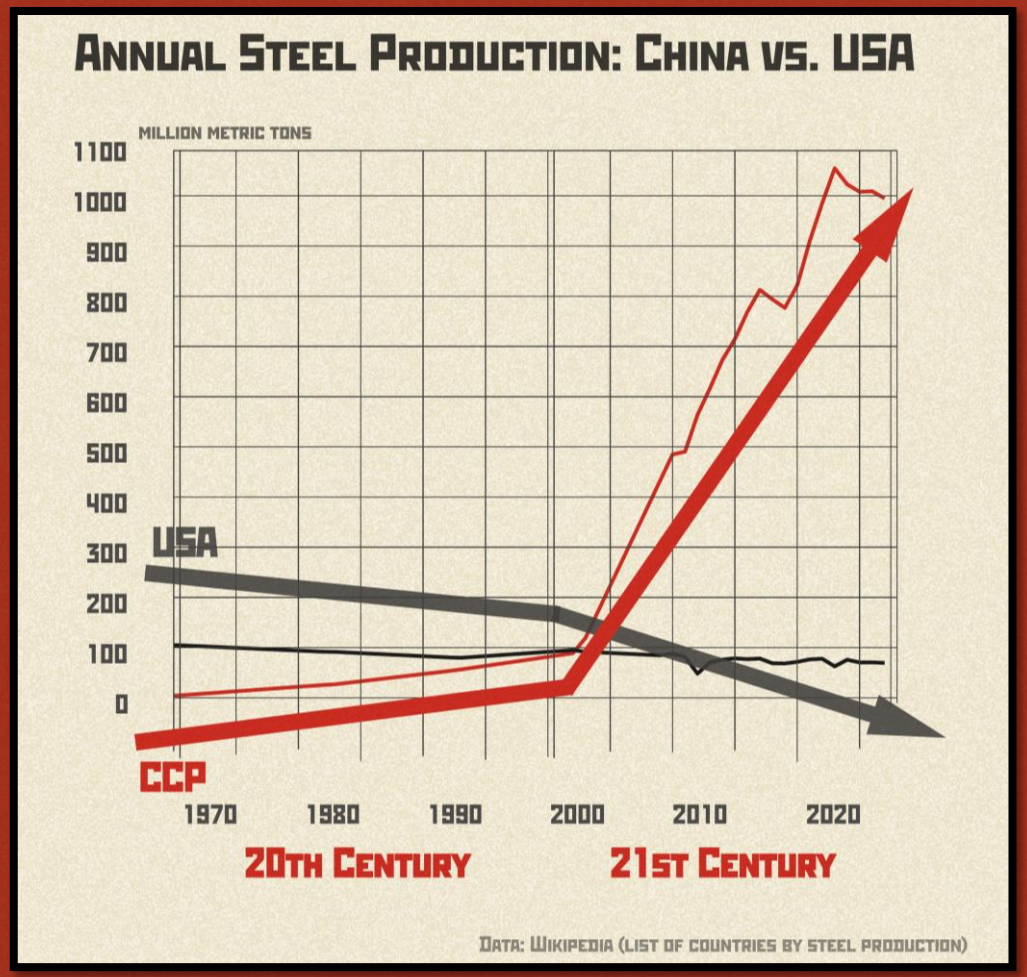
KServe 3

TorchServe 10

ONNX Runtime 11

The Fellowship of Kaggle

@DataCommissar



Visualization of Data in the Soviet Union


ДРОГНУЛА КАПИТАЛИСТИЧЕСКАЯ АМЕРИКА
MILLION METRIC
1100
1000
900
800
700
600
500
400
300
200
100
0
СССР
USA
1970
1980
1990
2000
2010
2020
20TH CENTURY
21ST CENTURY
DATA: WIKIPEDIA (LIST OF COUNTRIES BY STEEL PRODUCTION


ДРОГНУЛА КАПИТАЛИСТИЧЕСКАЯ АМЕРИКА
USA
1970
1980
1990
2000
2010
2020
20TH CENTURY
21ST CENTURY

I want to

go beyond using bar charts or repeating the same chart type

I want to

go beyond just visualizing data like the GDP per capita data

I want to

go beyond just using the default themes and styles

GETTING INSPIRED

#30DayChartChallenge

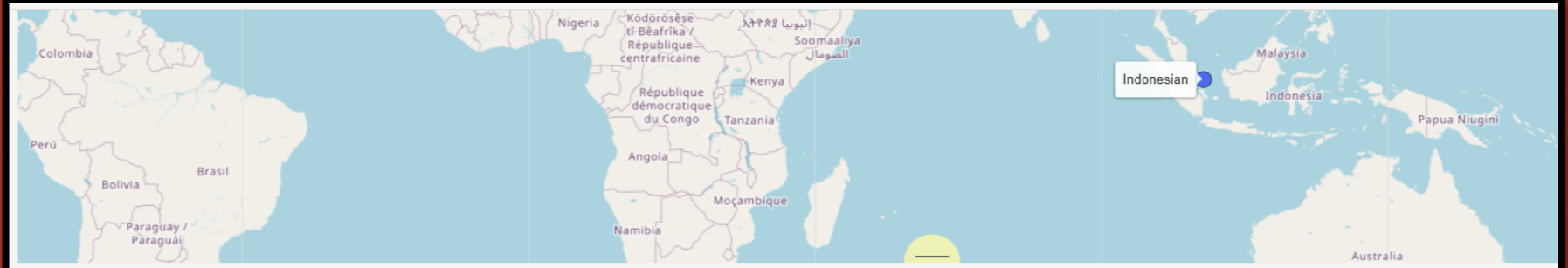
April 2022 • 30 Days • 30 Charts • 5 Categories

Comparisons	Distributions	Relationships	Timeseries	Uncertainties
1. part-to-whole	7. physical	13. correlation	19. global change	25. trend
2. pictogram	8. mountains	14. 3-dimensional	20. new tool	26. interactive
3. historical	9. statistics	15. multivariate	21. down/upwards	27. future
4. flora	10. experimental	16. environment	22. animation	28. deviations
5. slope	11. circular	17. connections	23. tiles	29. storytelling
6. data day: OWID	12. theme day: The Economist	18. data day: OECD	24. theme day: Financial Times	30. data day: UN Population

Follow @30DayChartChall for more!

#30DayChartChallenge 2022 Edition

Data: Dutch Loanwords



Showing 1 to 20 of 328 entries

← Previous 1 2 3 4 5 Next →

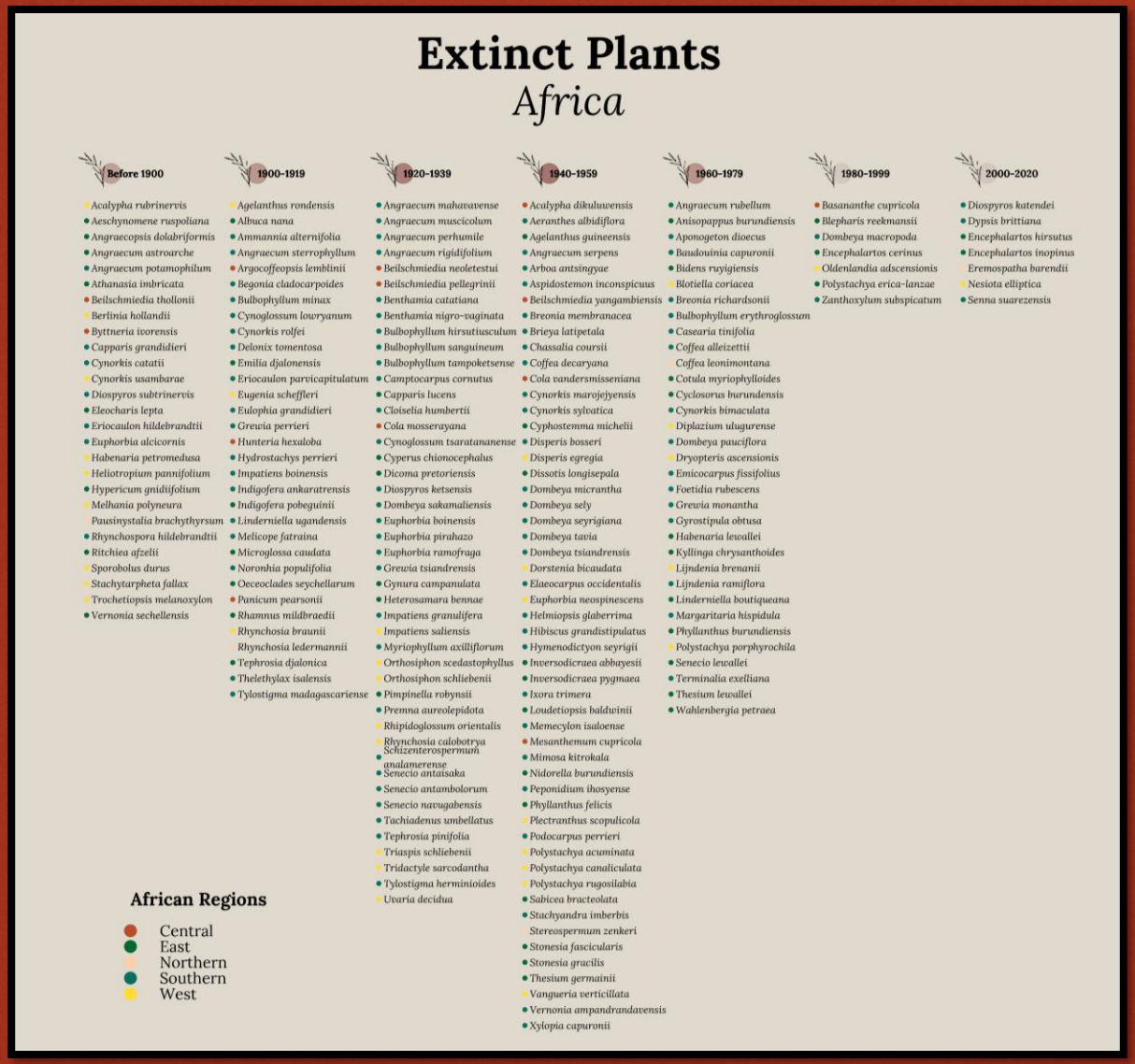
Borrowed words	Relation	Donor languoid	Source word
Search	--any--	Search	Search
aanbidden	immediate	Latin	*adorare*
adelaar	immediate	High German	*Adler*
adres	immediate	French	*adresse*
alligator	immediate	English	*alligator*
altaar	immediate	Latin	*altāre*
anker	immediate	Latin	*ancora*
arts	immediate	High German	*Artz*
auto	immediate	English	*auto(mobile)*
baai	immediate	French	*baie*
baarmoeder	immediate	Latin	*Latijn matrix*
baby	immediate	English	*baby*

World Loanword Database (WOLD)

Data: Dutch Loanwords

Dutch Word	Source Language	Source Word	Meaning
 kaas	Latin	cāseus	cheese
 kaars	High German	kerza	candle
 rijk	Celtic	riks	rich
 haai	Old Norse	hár	shark

Searching for text-heavy chart inspirations

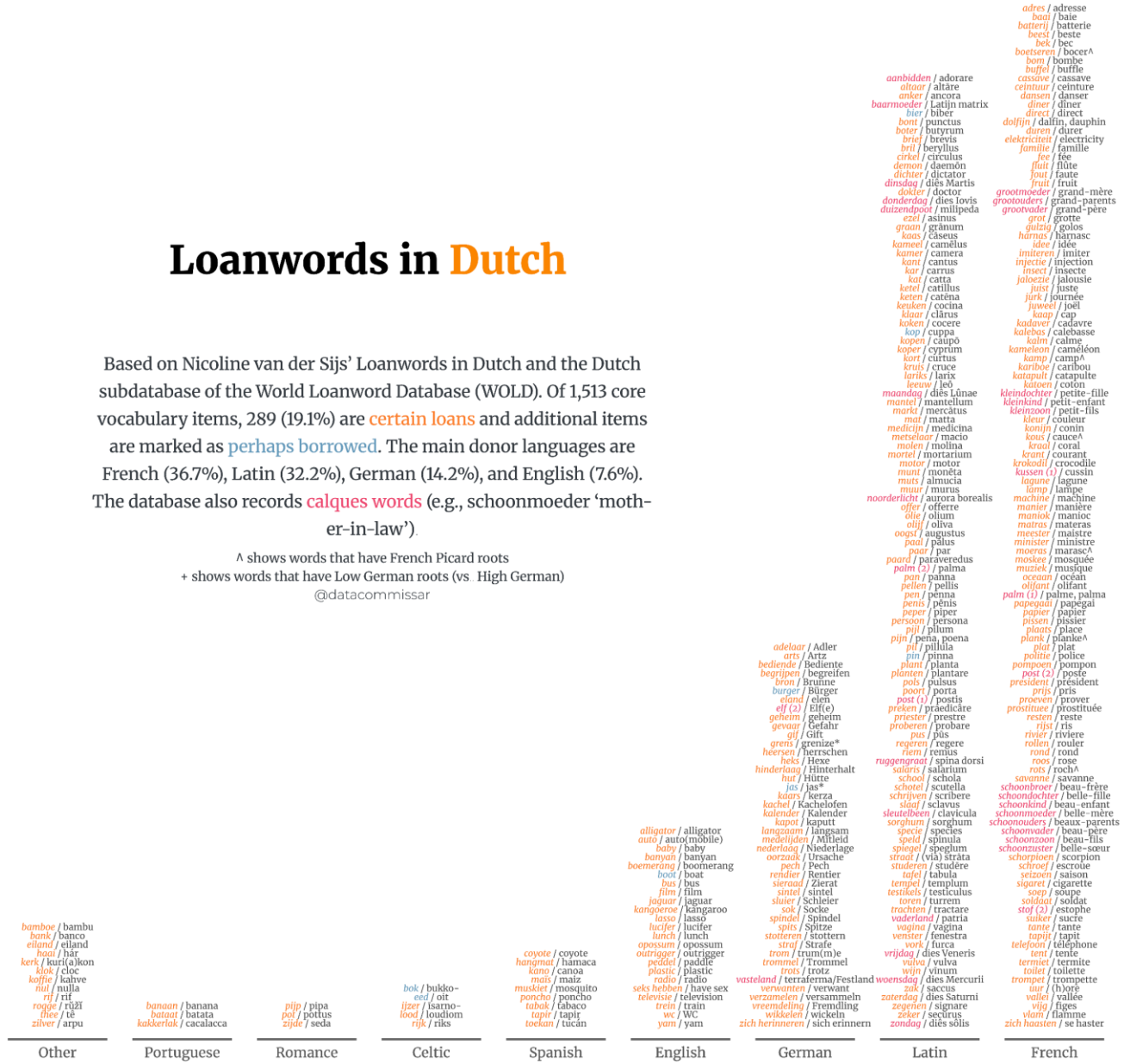


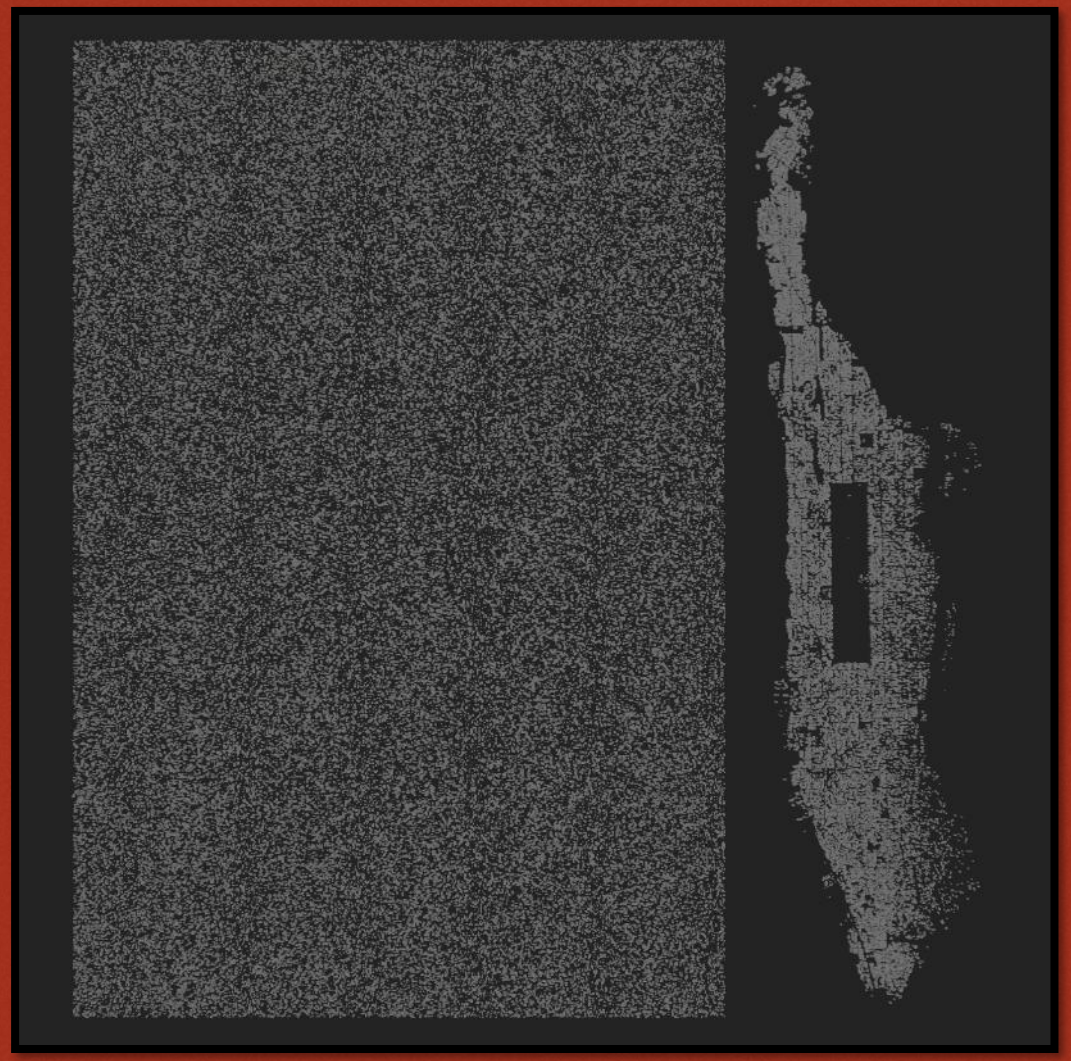
Ijeamaka Anyene, IUN declared extinct plants in Africa

Loanwords in Dutch

Based on Nicoline van der Sijs' Loanwords in Dutch and the Dutch subdatabase of the World Loanword Database (WOLD). Of 1,513 core vocabulary items, 289 (19.1%) are certain loans and additional items are marked as perhaps borrowed. The main donor languages are French (36.7%), Latin (32.2%), German (14.2%), and English (7.6%). The database also records calques words (e.g., schoonmoeder 'mother-in-law').

^ shows words that have French Picard roots
+ shows words that have Low German roots (vs. High German)
@datacommissar





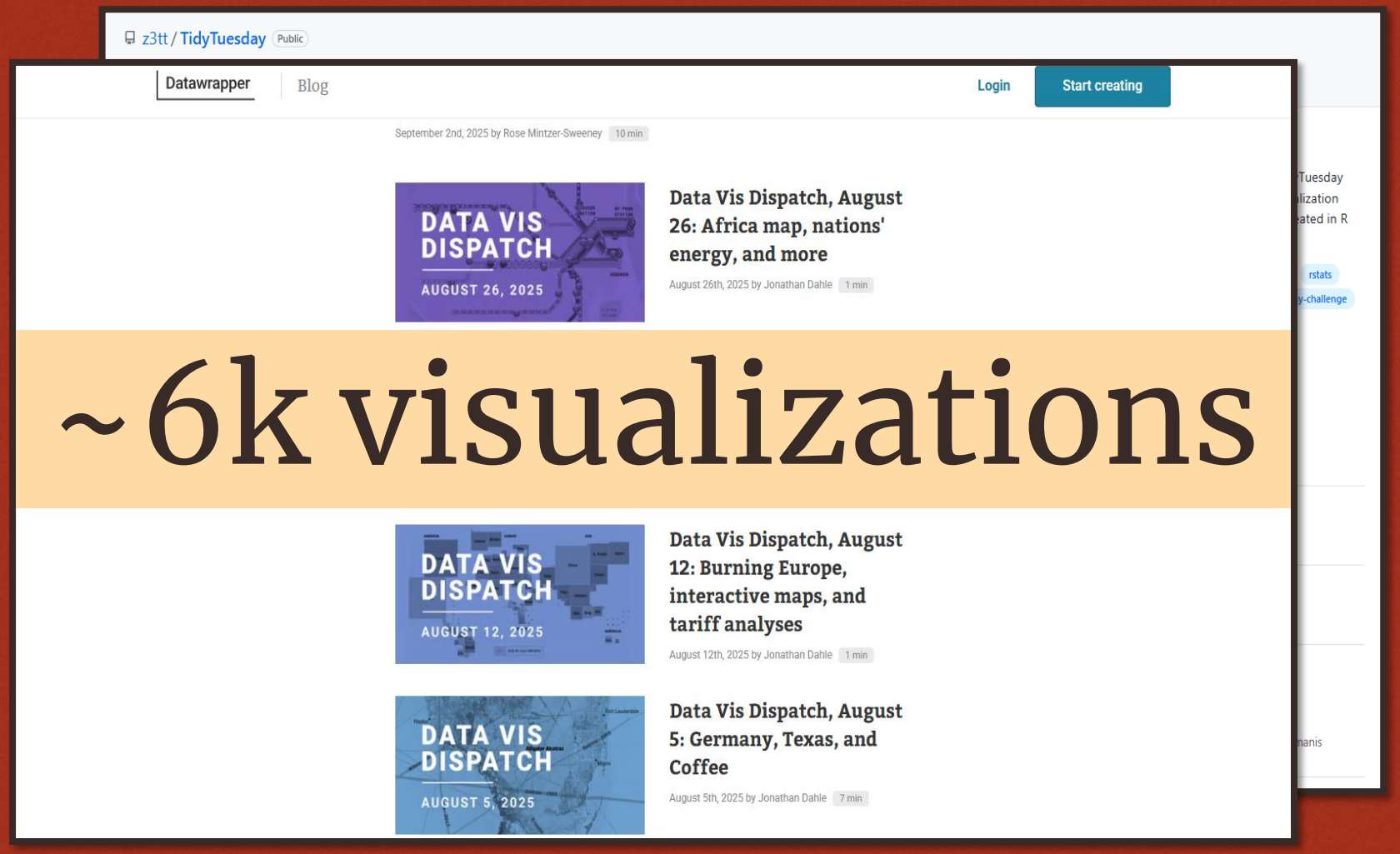
Reuters: Gaza in rubble and ruin

#TidyTuesday Contributions

charts + code

~2k contributions

Datawrapper Data Vis Dispatch



Flowingdata.com

z3tt / TidyTuesday Public

Datawrapper Blog Login Start creating

FLOWINGDATA Membership Books Projects Learning Newsletters Member Login

Data Underload / occupation, salary, work

Salary and Occupation

The median salary for full-time workers in the United States was $49,500, based on estimates from the Bureau of Labor Statistics in 2024. However, salaries vary by occupation. These charts show the spread.

Support an independent FlowingData, and get extra visualization goodness.

BECOME A MEMBER →

August 21, 2025
Visualization editing

August 14, 2025
Careless chart mistakes

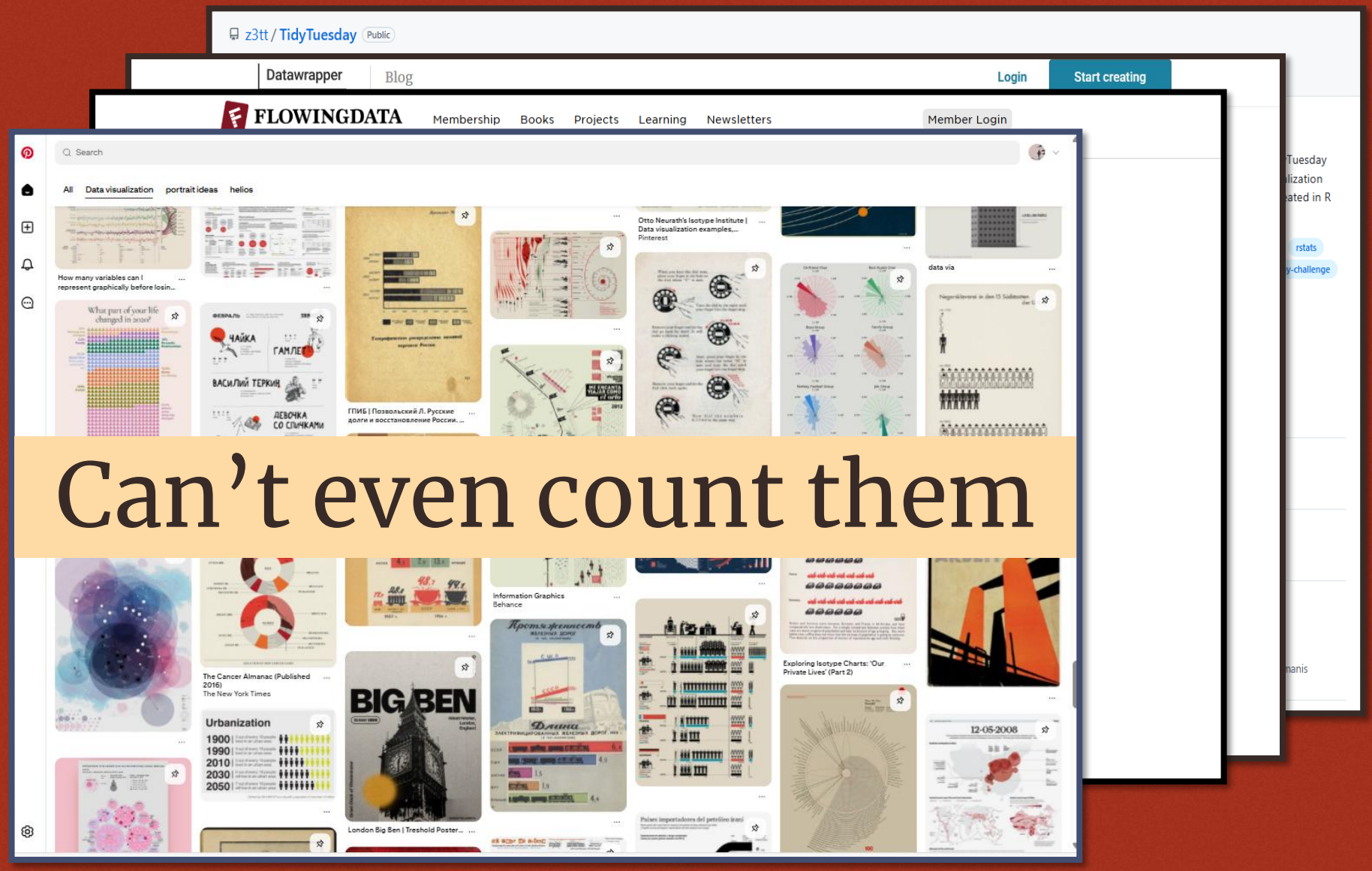
August 7, 2025
Familiar chart advantages

Second Edition

Visualize This: The FlowingData Guide to Design, Visualization, and Statistics (2nd Edition)
New tools, refined process.
Order: Amazon / Bookshop

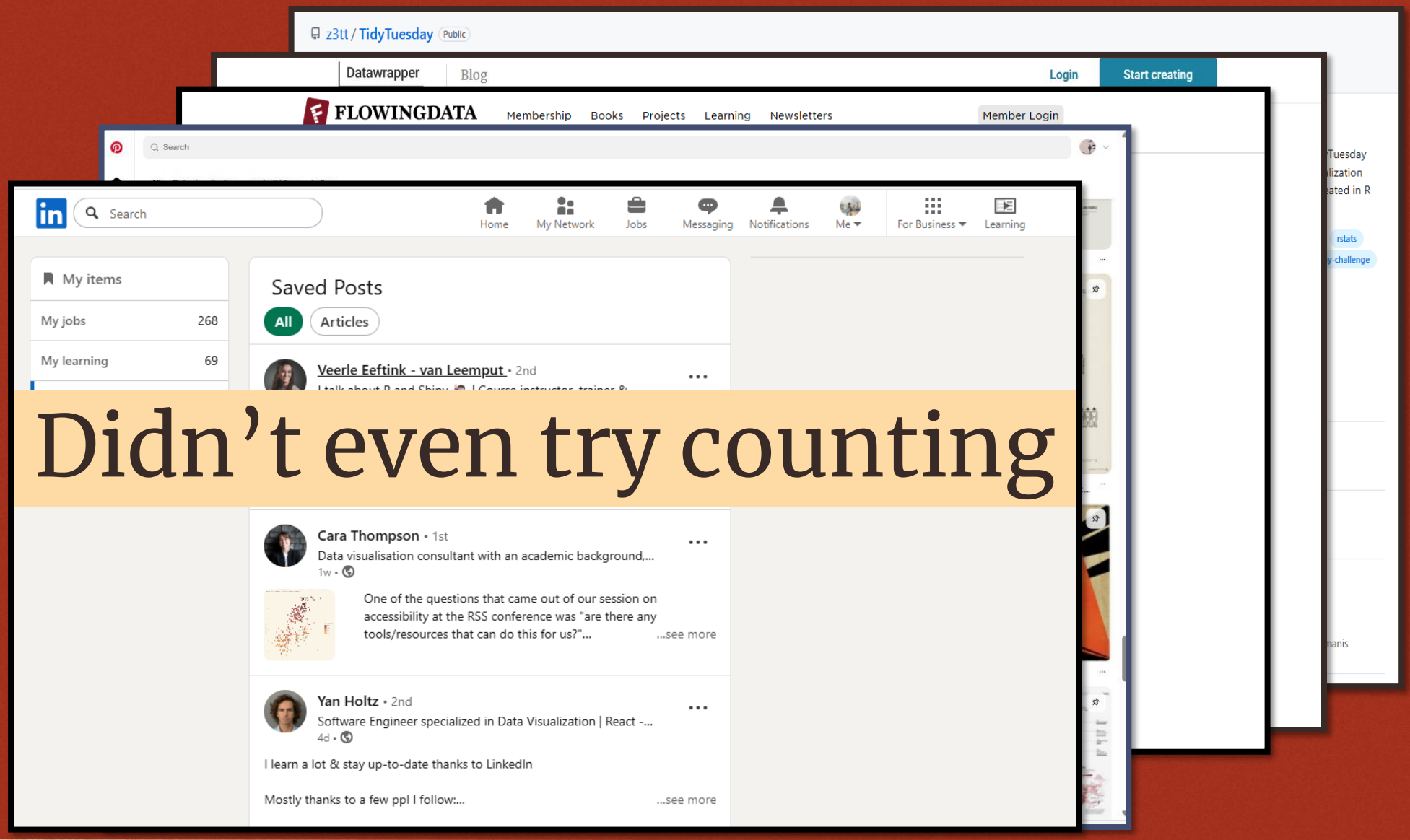
~9k visualizations

Pinterest

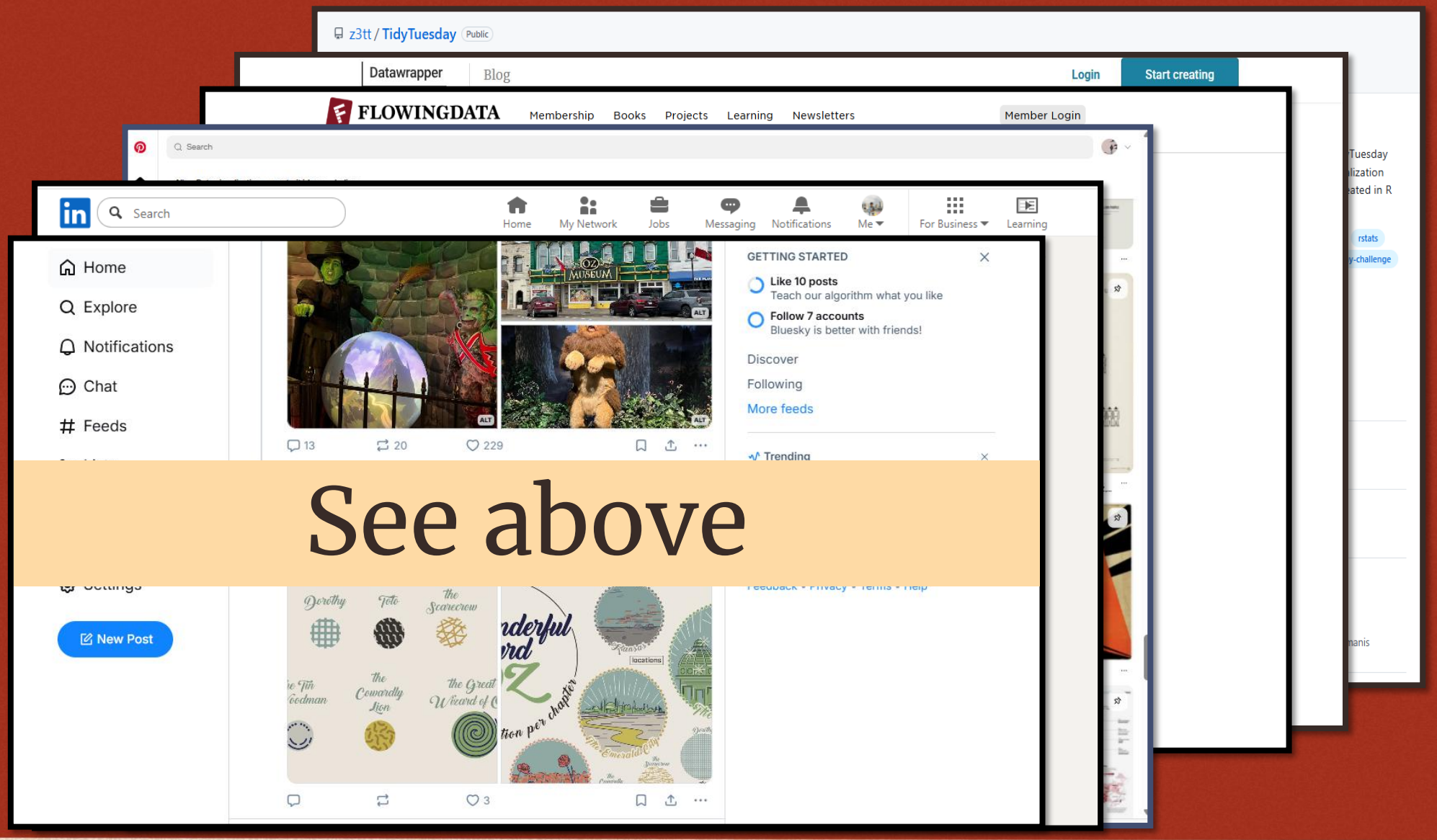


Can't even count them

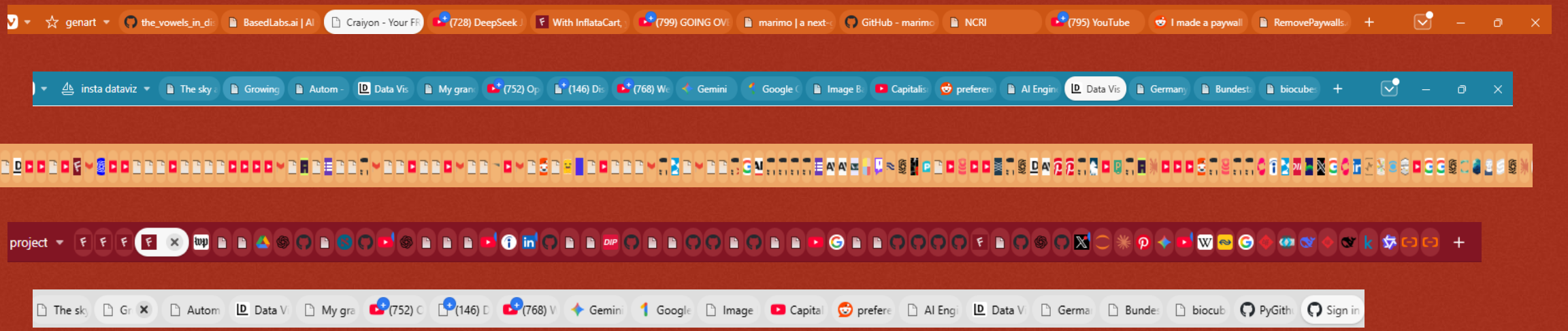
LinkedIn


z3tt / TidyTuesday Public
Datawrapper
Blog
Login
Start creating
FLOWINGDATA
Membership
Books
Projects
Learning
Newsletters
Member Login
Search
Home
My Network
Jobs
Messaging
Notifications
Me
For Business
Learning
My items
My jobs 268
My learning 69
Saved Posts
All
Articles
Veerle Eeftink - van Leemput • 2nd
Didn't even try counting
Cara Thompson • 1st
Data visualisation consultant with an academic background,...
1w •
One of the questions that came out of our session on accessibility at the RSS conference was "are there any tools/resources that can do this for us?"...
...see more
Yan Holtz • 2nd
Software Engineer specialized in Data Visualization | React -...
4d •
I learn a lot & stay up-to-date thanks to LinkedIn
Mostly thanks to a few ppl I follow:...
...see more

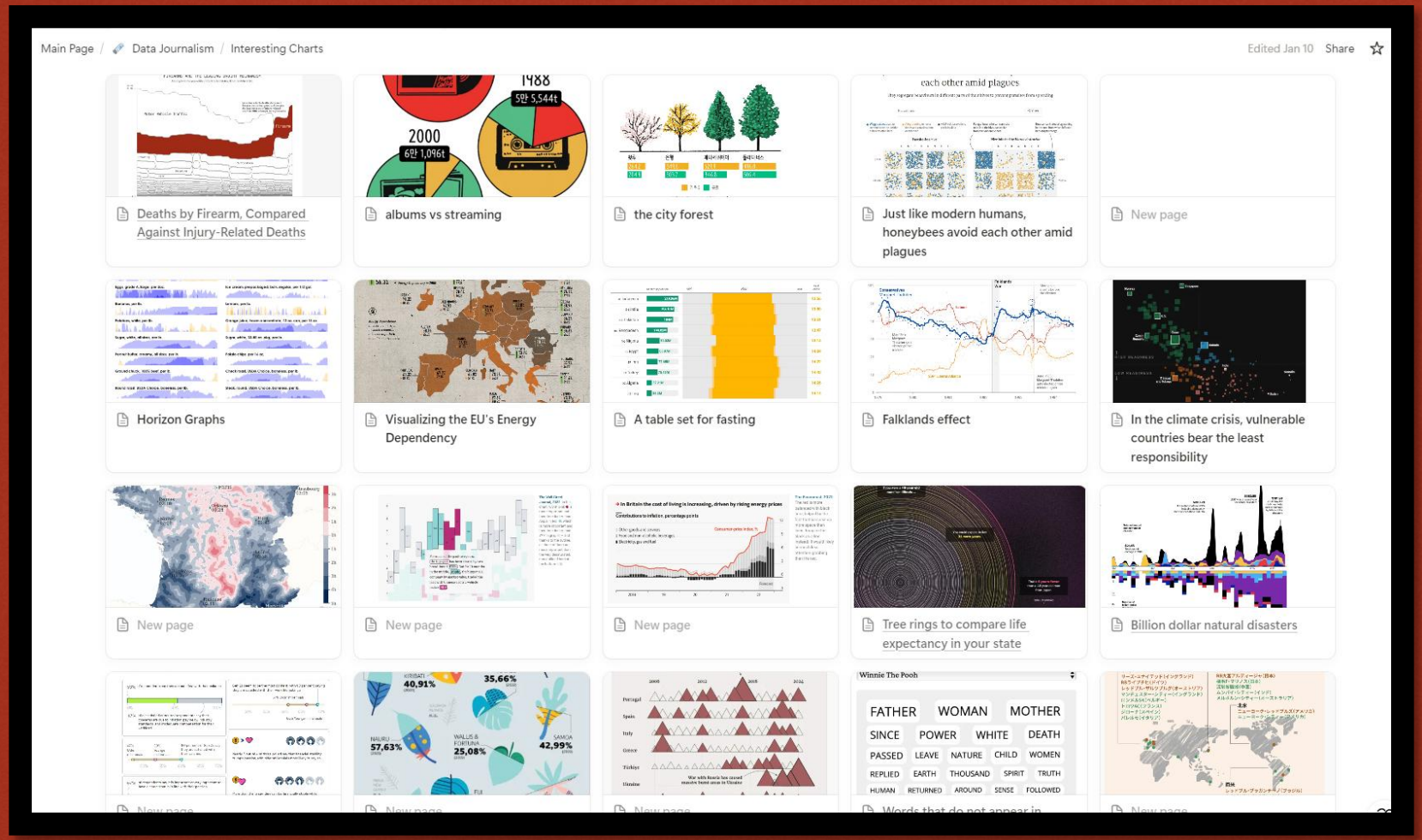
Bluesky

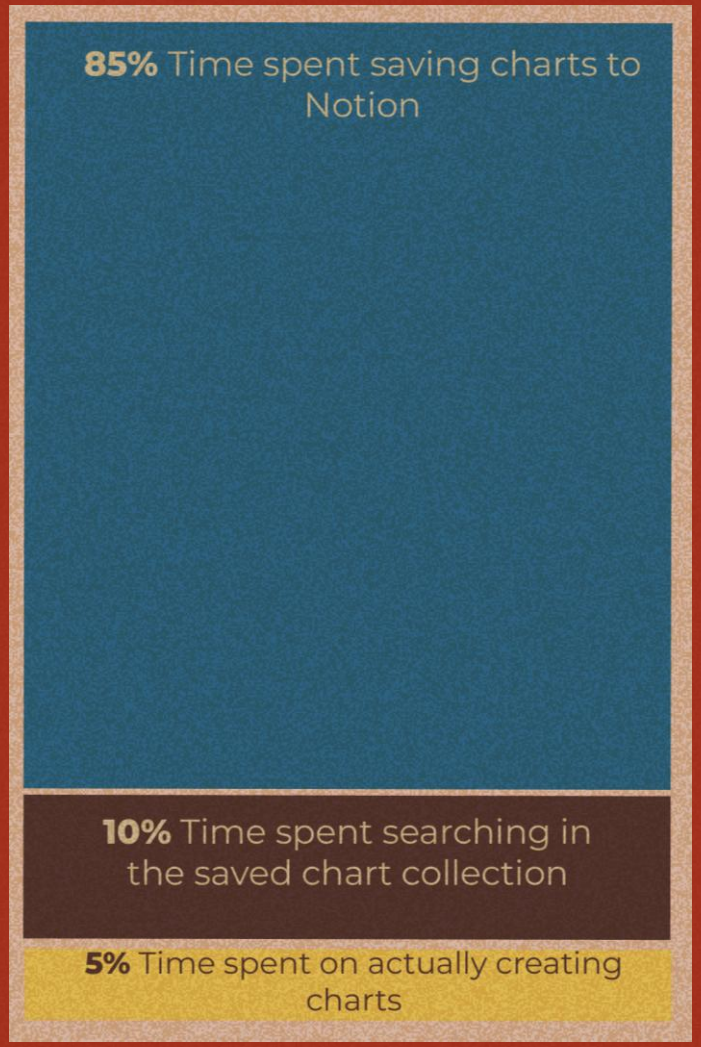


See above

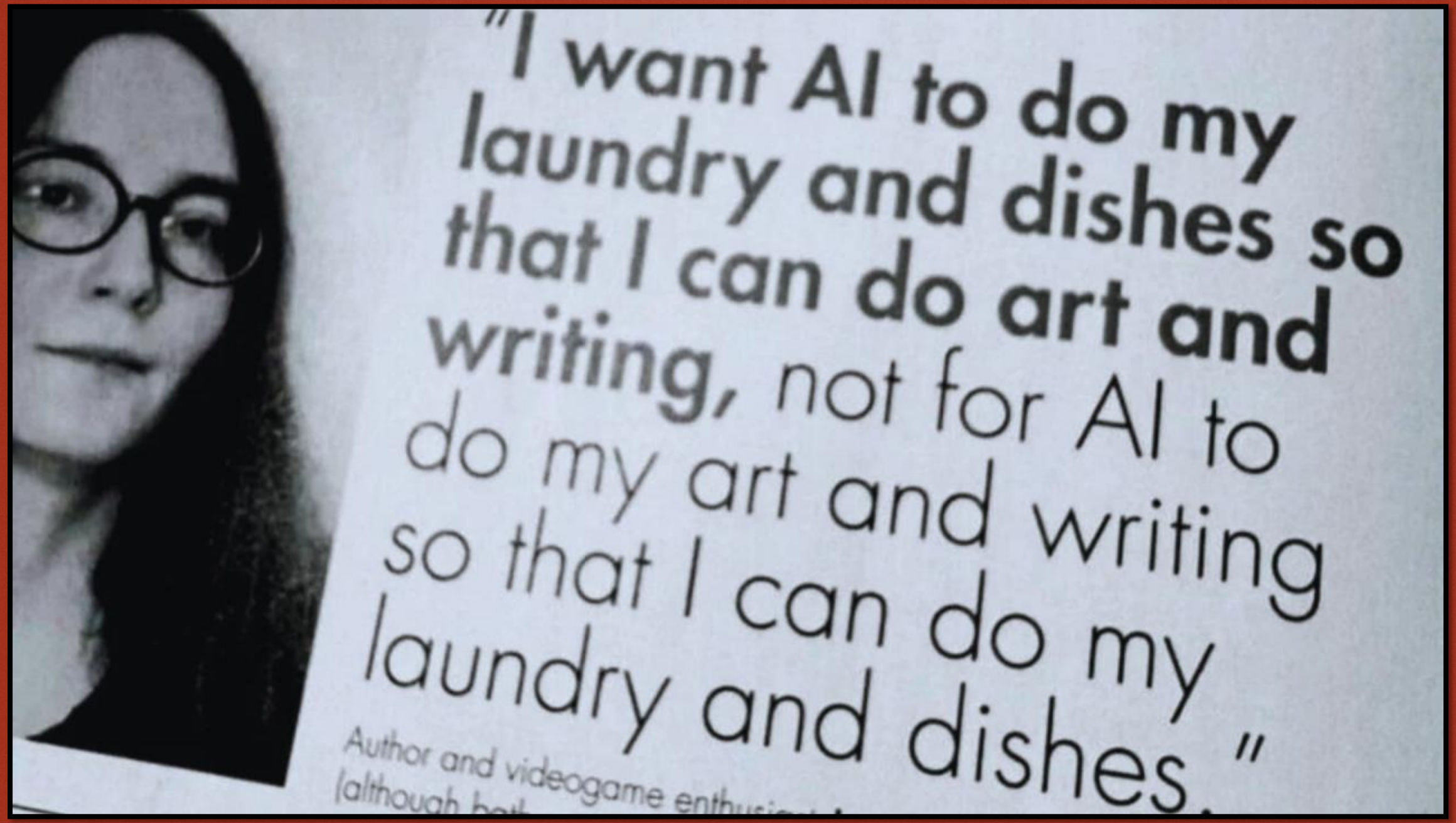

genart
Craiyon - Your Fr
(728) DeepSeek
With InflataCart,
(799) GOING OV
marimo | a next-
GitHub - marimo
NCRI
(795) YouTube
I made a paywall
RemovePaywalls
insta dataviz
The sky
Growing
Autom -
Data Vis
My gran
(752) Op
(146) Dis
(768) We
Gemini
Google
Image Ba
Capitalis
preferen
AI Engin
Data Vis
Germany
Bundest
biocubes
project
The sky
Gr
Autom
Data V
My gra
(752) C
(146) D
(768) V
Gemini
Google
Image
Capital
prefere
AI Engi
Data V
Germa
Bundes
biocub
PyGith
Sign in

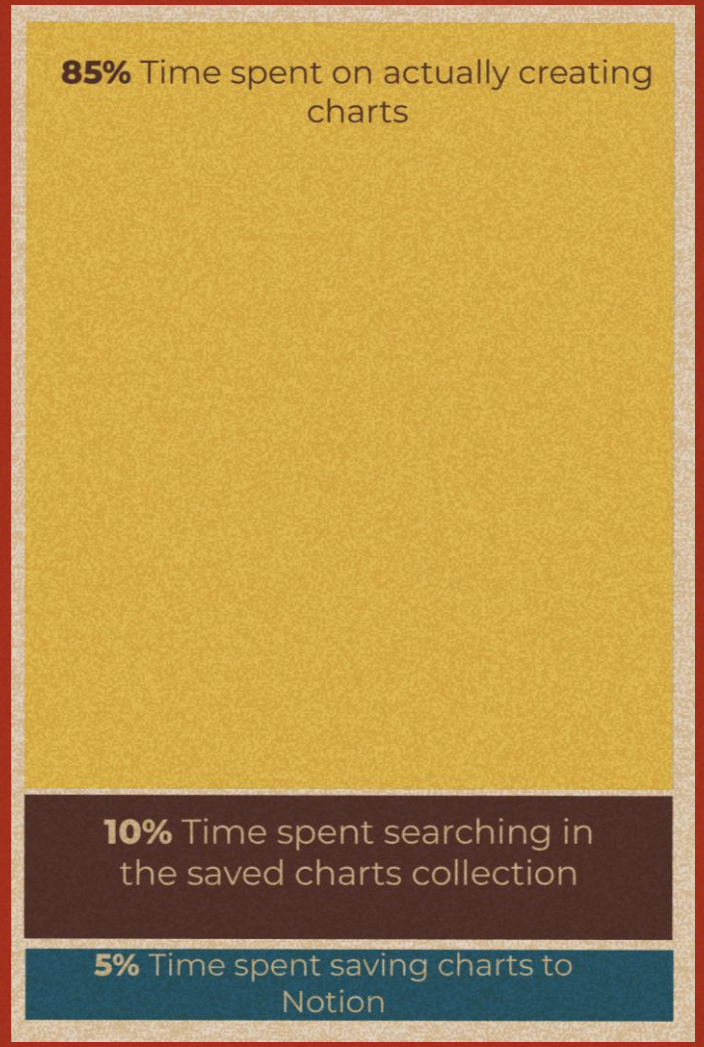
Solution 3: NOTION




85% Time spent saving charts to Notion
10% Time spent searching in the saved chart collection
5% Time spent on actually creating charts


Newsletters, twitter, LinkedIn ...
Me trying to keep up with new visualizations
1,280


"I want AI to do my laundry and dishes so that I can do art and writing, not for AI to do my art and writing so that I can do my laundry and dishes."
Author and videogame enthusi
(although both


85% Time spent on actually creating charts
10% Time spent searching in the saved charts collection
5% Time spent saving charts to Notion

What if, instead of scrolling, I built a system that can semantically retrieve charts?

Examples

I’m looking for charts house ownership rates

I’m looking for charts that show proportions or distributions

I’m looking for circular bar plots

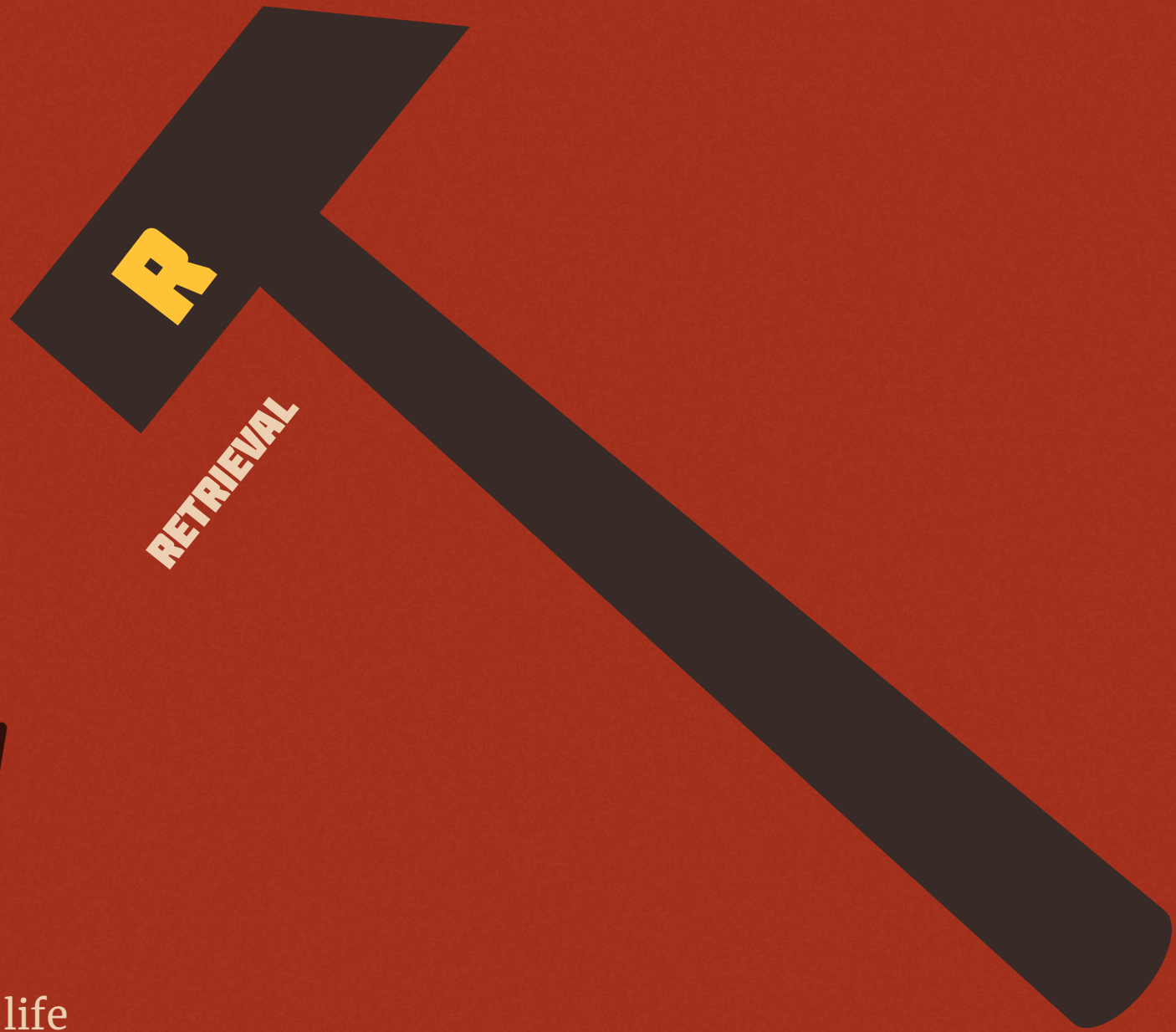
I’m looking for charts that have diverging legends

problems in my life

IT'S DEMO TIME

GO TO
MYNEXTCHART.COM

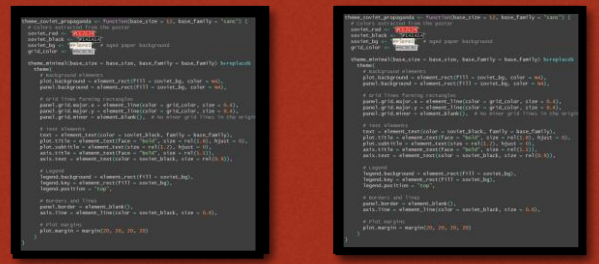
TECHNICAL IMPLEMENTATION



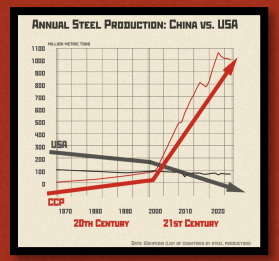
problems in my life

Retrieving Charts

ggplot2 code



Plots



Embedding Model

my query

Vector Database

charts

Embedding Model

my query

Vector Database

charts

From Graphs to Vectors

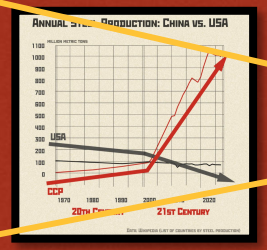
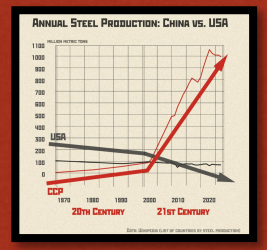


IMAGE EMBEDDINGS



VLM

PLOT DESCRIPTION

TEXT EMBEDDINGS

Misses Details

TEXT EMBEDDINGS

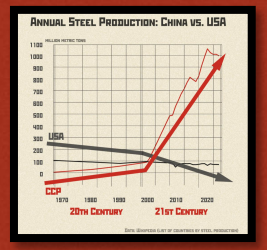
LLM

PLOT DESCRIPTION

TEXT EMBEDDINGS

Misses the big picture (literally)

From Graphs to Vectors



Only plots made by ggplot2 from Github

VLM

QWEN-VL-MAX

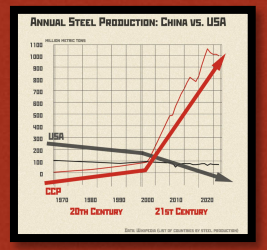
PLOT DESCRIPTION

TEXT EMBEDDINGS

MISTRAL AI_

Mistral Embeddings

From Graphs to Vectors



Only plots made by ggplot2 from Github

VLM

QWEN-VL-MAX

prompt

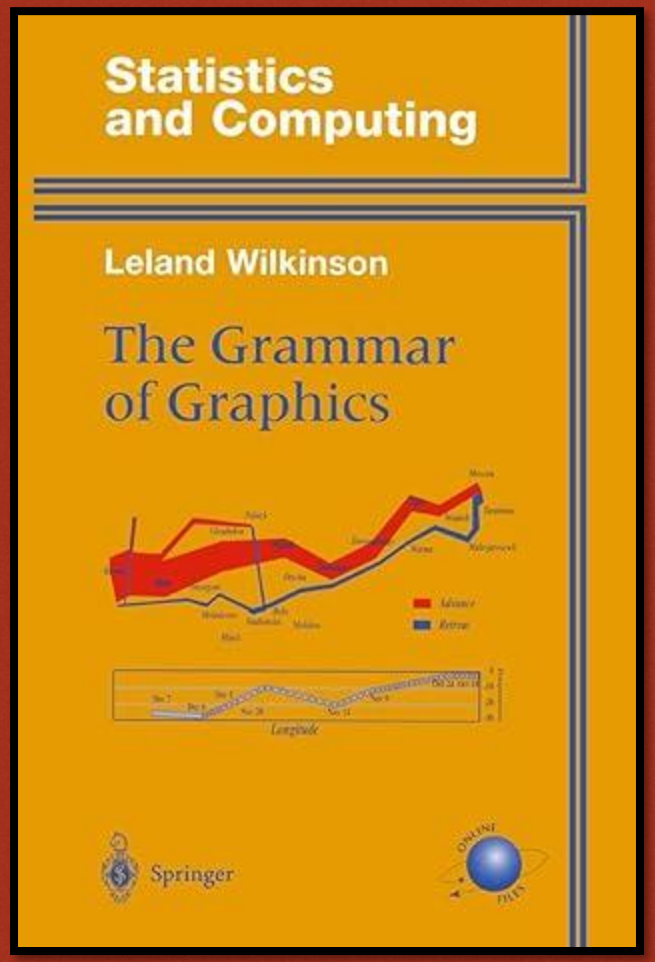
PLOT DESCRIPTION

TEXT EMBEDDINGS

MISTRAL AI_

Mistral Embeddings

Grammar of Graphics

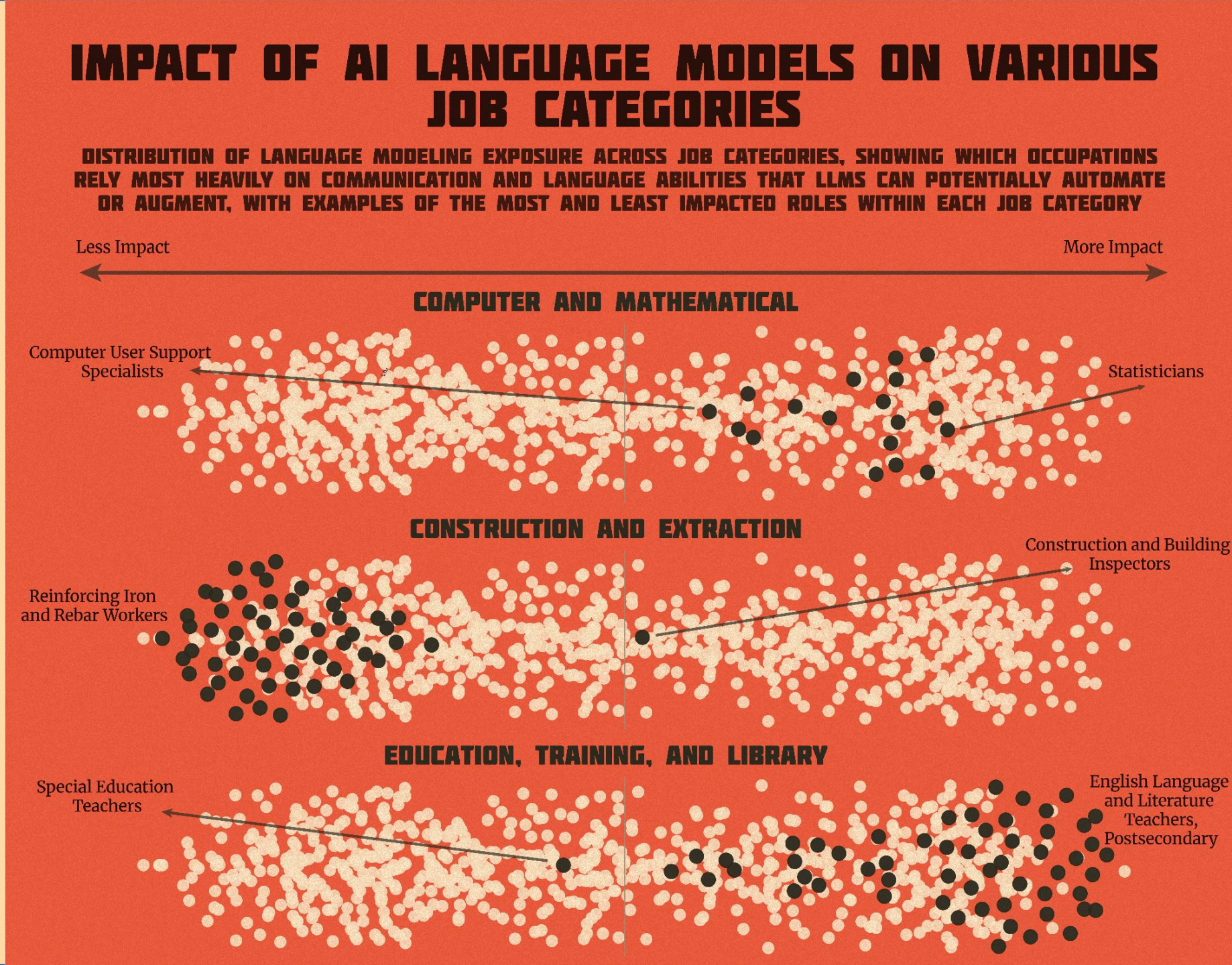


Data Visualization by Andrew Heiss

Grammar of Graphics

Component	ggplot2 (R)	plotnine (Python)	Explanation
Data	ggplot(data)	ggplot(data)	The raw data that you want to visualize (initializing a plot).
Aesthetics	aes()	aes()	The mapping between variables and visual properties.
Geometries	geom_*()	geom_*()	The geometric shape of a layer representing the data.
Statistics	stat_*()	stat_*()	The statistical transformation of a layer applied to the data.
Scales	scale_*()	scale_*()	The representation of mapped aesthetic attributes.
Coordinate System	coord_*()	coord_*()	The transformation to map data coordinates into the plot plane.
Facets	facet_*()	facet_*()	The arrangement of the data into a set of small multiples.
** Themes*	*theme()*	*theme()*	*The overall visual defaults of non-data elements of the graphic.*

Lets' look at the theme



Theme

Background color: #E95A3C

Title: Center aligned

Subtitle: Center aligned

Title font: **MOLOT**

Subtitle font: **MOLOT**

Font color (black),

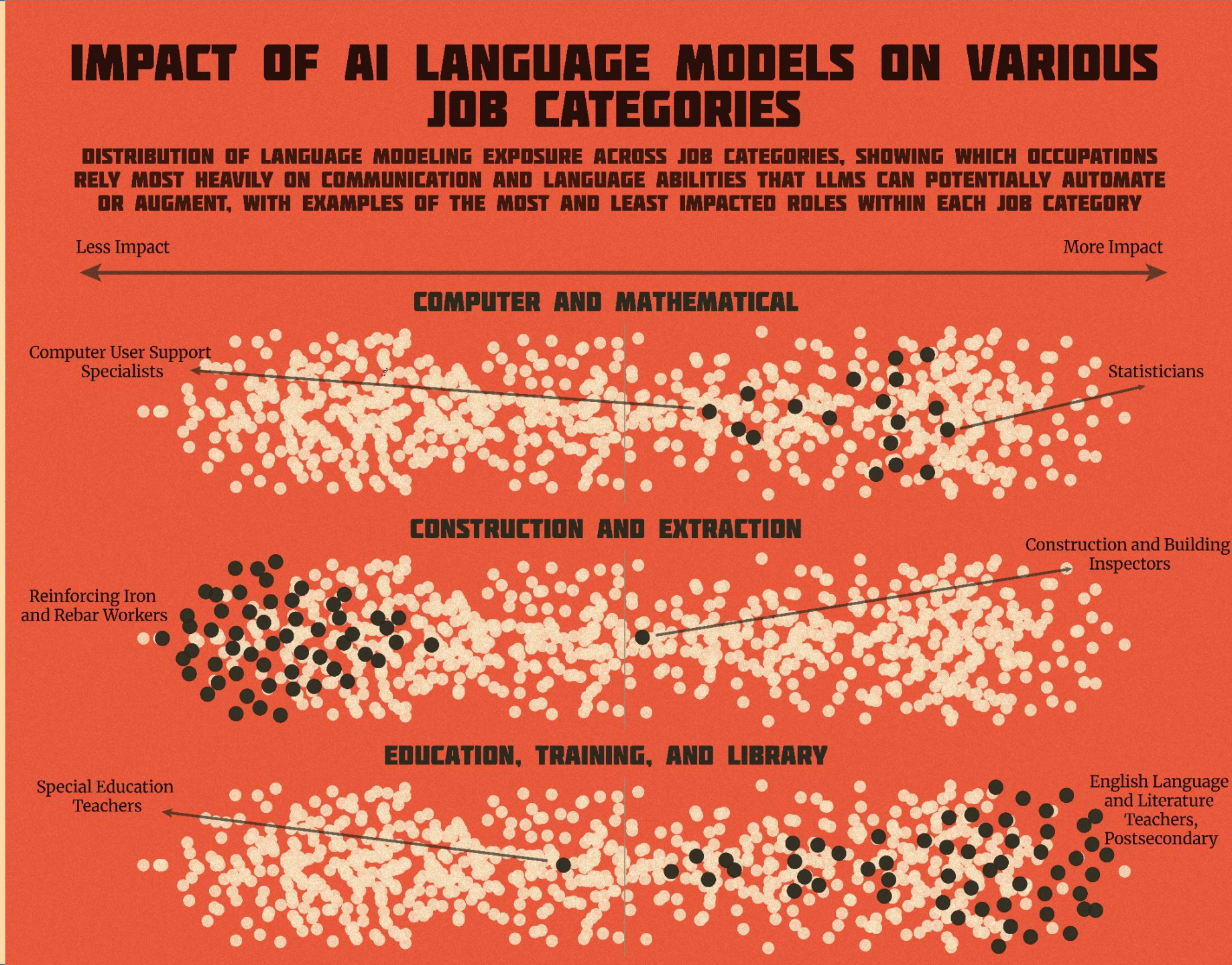
Font style (**bold**)

Legend: **None**

Grids: **None**

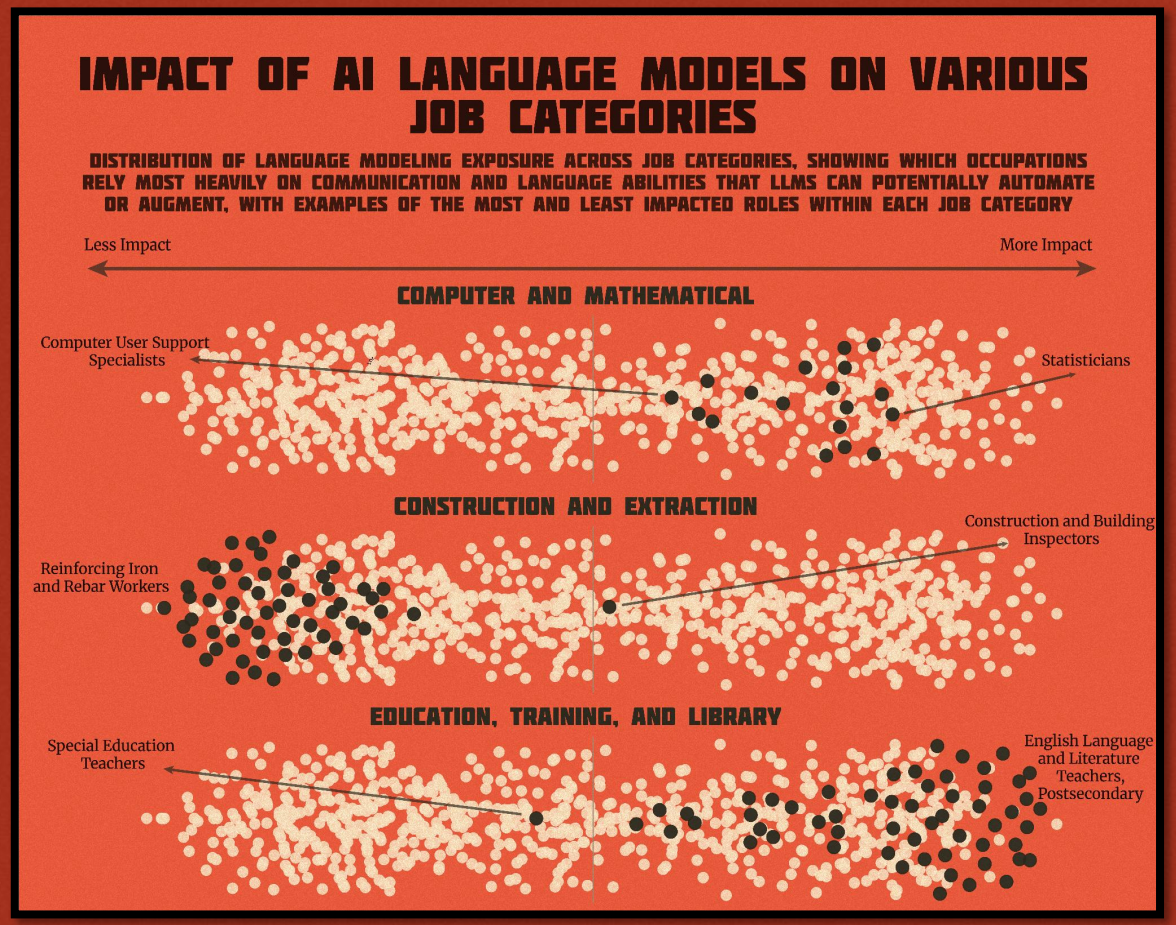
Axis: **None**

etc.



prompt_v1.txt

```
You're an expert in analyzing ggplot2 visualizations and R code. Your goal is to provide a comprehensive technical breakdown of the following R code that uses ggplot2 to create a visualization.
Format your response with these numbered sections separated by %%%:
0. **Plot Category & Purpose**
   - Understand and describe what this plot represents in high-level terms, for example: "This code creates a visualization showing the distance between each US state's capital city and its largest city.
   - Describe the type of data being visualized and the overall purpose of the plot.
   - Describe the visualization approach in detail, including the type of plot, the data being used, and any specific features or patterns that stand out.
   - Target audience analysis (who would benefit most from this visualization)
   - Visualization purpose (exploratory vs explanatory)
   - Alternative visualization approaches that could present the same data
   - The logic of the data visualization
   - Based on the categories of the previous editions of #30DayChartChallenge, identify which categories apply (separated by commas): Experimental, Pop Culture, Stats & Models, Comparison, Distribution,

1. **Data Source & Preparation**
   - Data source identification and retrieval methods
   - Key data transformations before plotting
   - Data cleaning and filtering techniques applied
   - Variable creation and manipulation
   - External data merges or joins

2. **Title & Text Elements**
   - Plot title, subtitle, and caption with exact text
   - Annotations and direct labels (geom_text/label, annotate()) with positioning parameters
   - Text formatting specifications (size, color, face)
   - Contextual information provided through text

3. **Geometry & Mappings**
   - All geometry types used (list each geom with full parameters)
   - Variable mappings for each geometry (what maps to x, y, color, fill, size, etc.)
   - Layer order and drawing precedence
   - Aesthetic mapping inheritance patterns

4. **Statistical Transformations & Coordinates**
   - Faceting implementation (facet_wrap/grid variables and scales)
   - Coordinate system details (coord_flip, coord_polar, limits, expansions)
   - Statistical transformations (stat_summary, binning, smoothing parameters)
   - Scale transformations (log, sqrt, etc.)

5. **Theme Architecture**
   - Base theme template (theme_minimal, theme_bw, etc.)
   - Complete theme() chain with exact parameter values
   - Notable theme overrides categorized by element type (text, line, rect)
   - Custom margin and spacing adjustments
   - Overall aesthetic intention of the theme choices

6. **Color Specifications**
   - Named color palette or custom scale functions with hex codes
   - Named background color and fill with hex codes
   - Manual color/fill values with hex codes
   - Transparency settings (alpha values)
```



%%%

0. **Plot Category & Purpose**

- **High-level description**:

  This visualization shows how **AI language models (LLMs)** are expected to impact different job categories. Each dot represents an occupation within that category, positioned along a horizontal axis from "Less Impact" to "More Impact." Specific occupations are highlighted and annotated to illustrate extremes (least and most affected).

- **Type of data**:

  Occupational data, impact score (continuous), categorical job groups, and highlight markers.

- **Visualization approach**:

  A **faceted strip plot / dot scatter plot**:

  - X-axis = degree of LLM impact.
  - Y-axis = different job categories arranged in stacked panels.
  - Cream dots = all occupations.
  - Black dots = highlighted specific jobs (annotated with arrows and labels).

3. **Geometry & Mappings**

- **Geoms**:

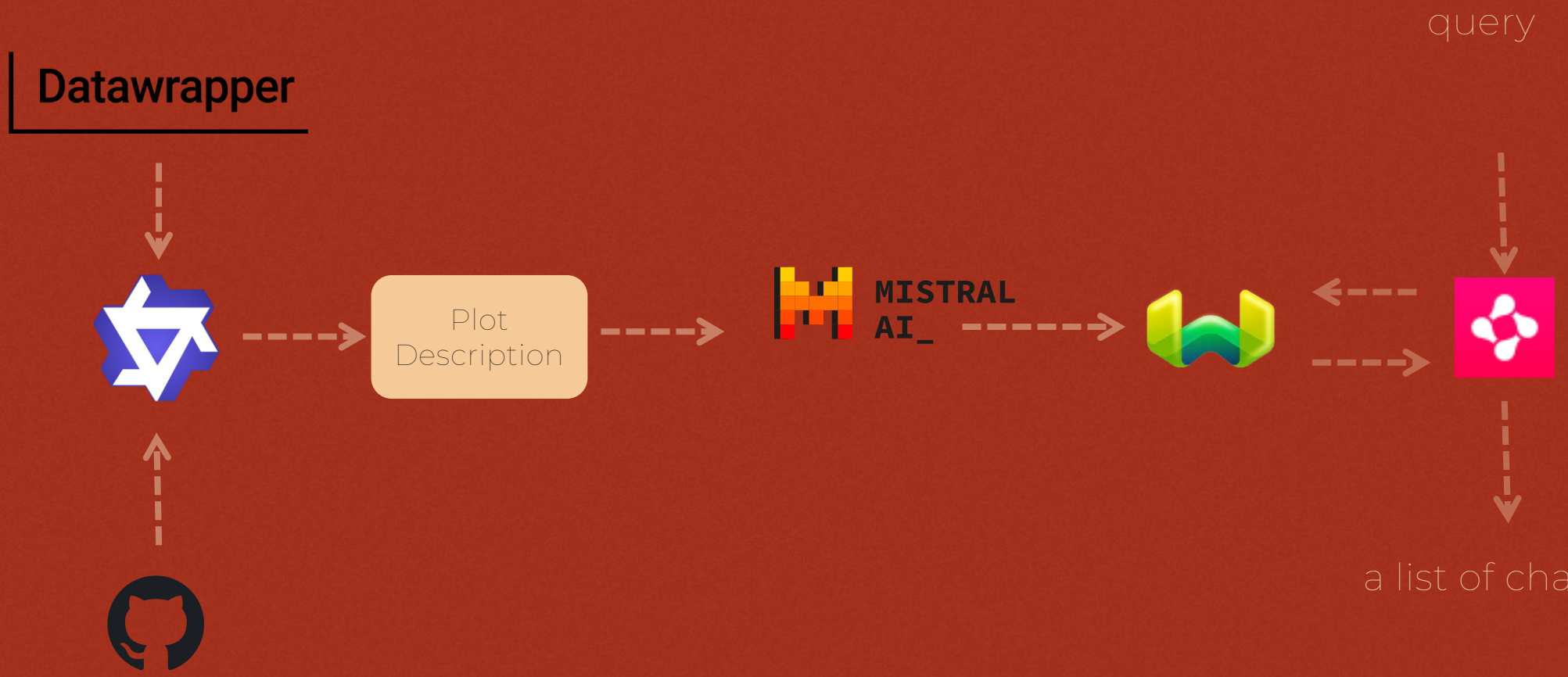
  - `geom_point(alpha=0.8, size=2)` → scatter of occupations.
  - `geom_point(size=2.5, color="black")` → highlight jobs.
  - `geom_text(hjust/vjust adjustments)` → occupation annotations.
  - `geom_segment(arrow=arrow(), linetype="solid")` → connecting arrows.

- **Mappings**:

MODE 1

VANILLA/HYBRID RETRIEVAL

THE PROBLEM WITH VANILLA RETRIEVAL



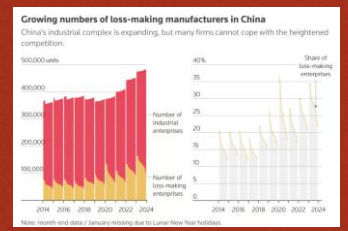
THE PROBLEM WITH VANILLA RETRIEVAL

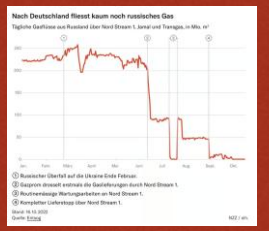
Datawrapper

Plot Description

MISTRAL AI_

I'm looking for a line chart with a red background color





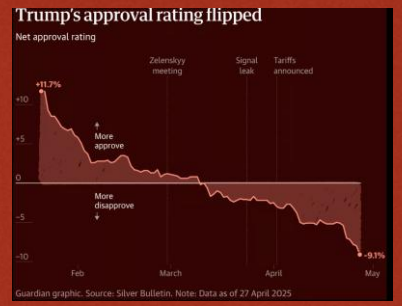
THE PROBLEM WITH VANILLA RETRIEVAL

Datawrapper

Plot Description

MISTRAL AI_

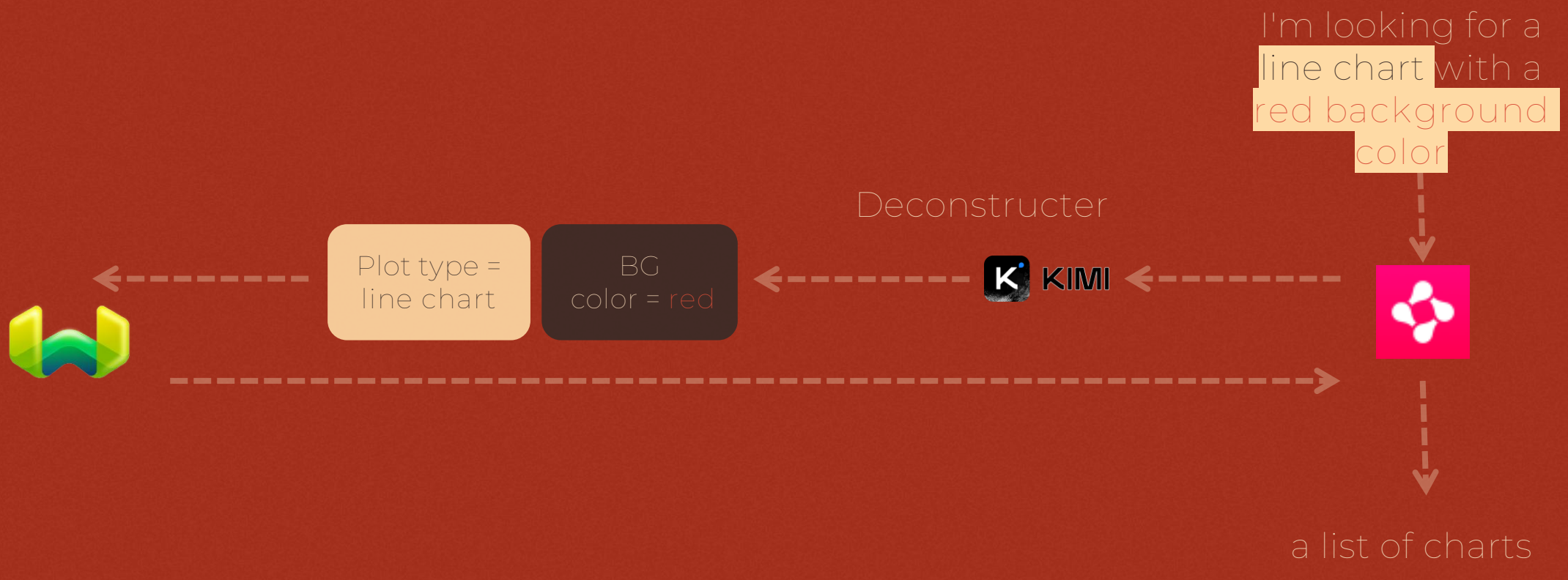
I'm looking for a line chart with a red background color



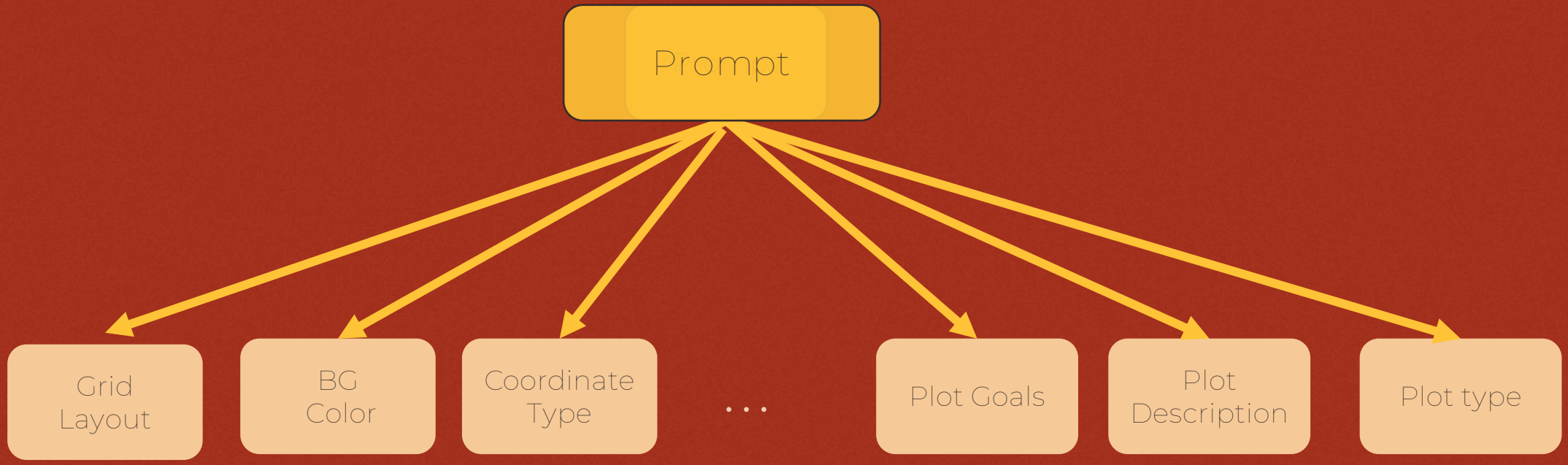
The “Semantic Soup” Problem

MODE 2

DECONSTRUCTING VISUALIZATIONS WITH TOOL CALLING

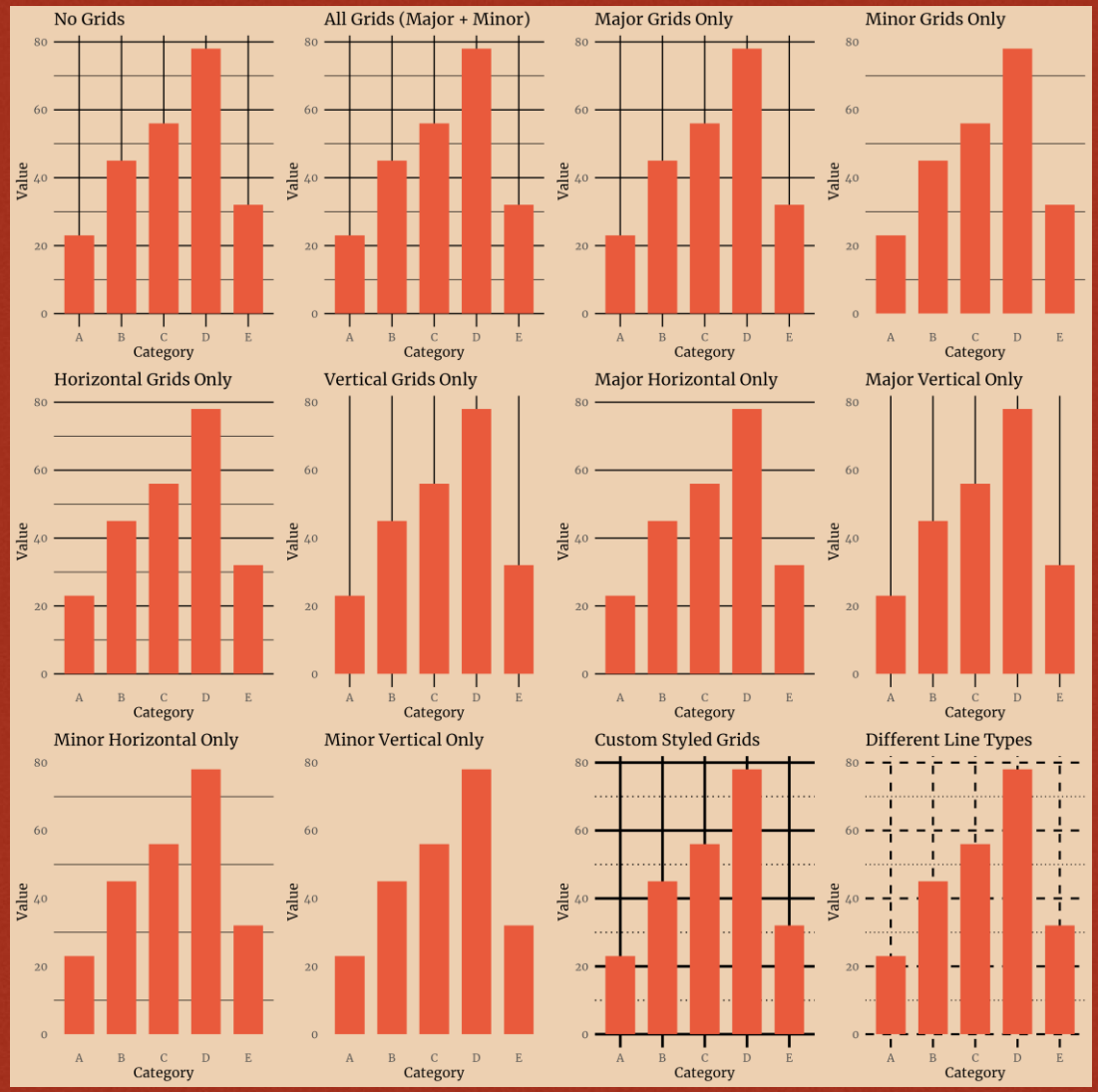

I'm looking for a line chart with a red background color
Deconstructer
KIMI
Plot type = line chart
BG color = red
a list of charts

Deconstructing the Prompt

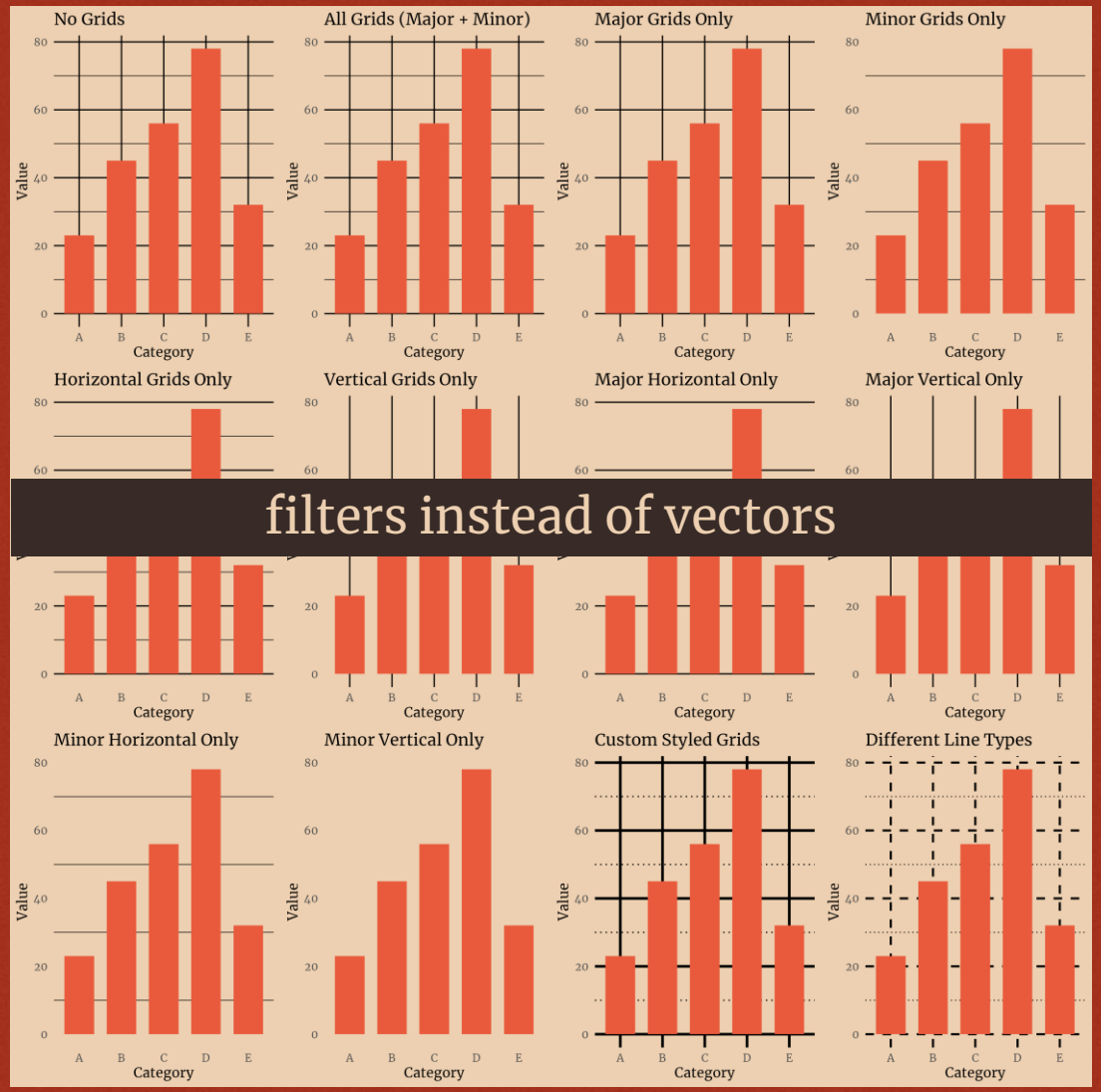


Each element stored as a vectorized field in the vector DB

Some chart elements have only a limited set of possible values.

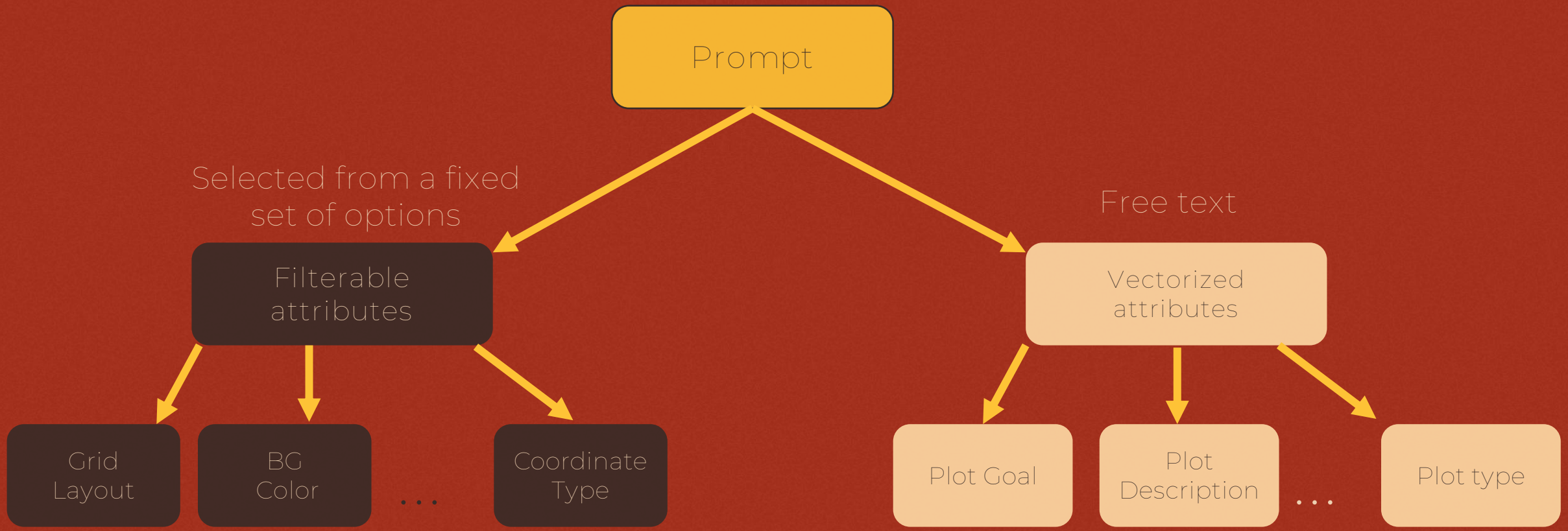


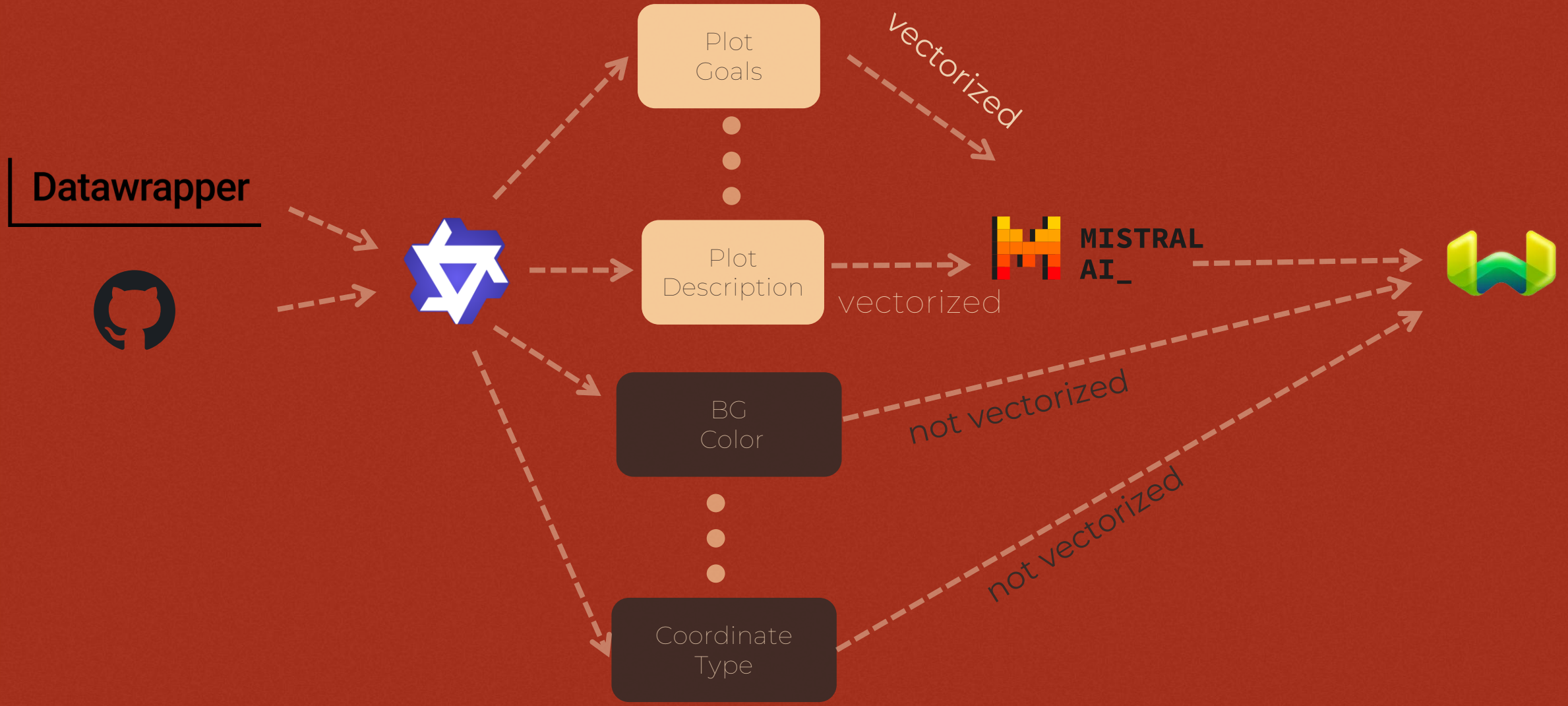
Some chart elements have only a limited set of possible values.



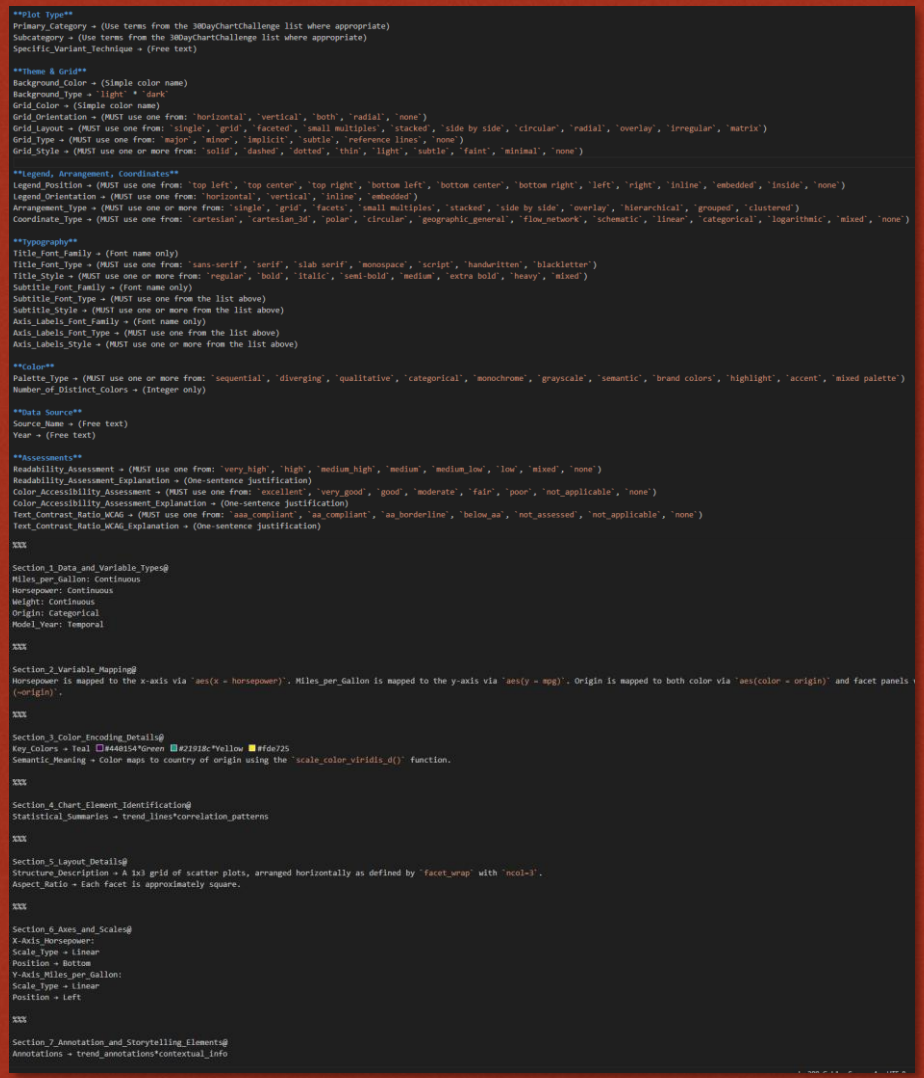
filters instead of vectors

Deconstructing the Prompt

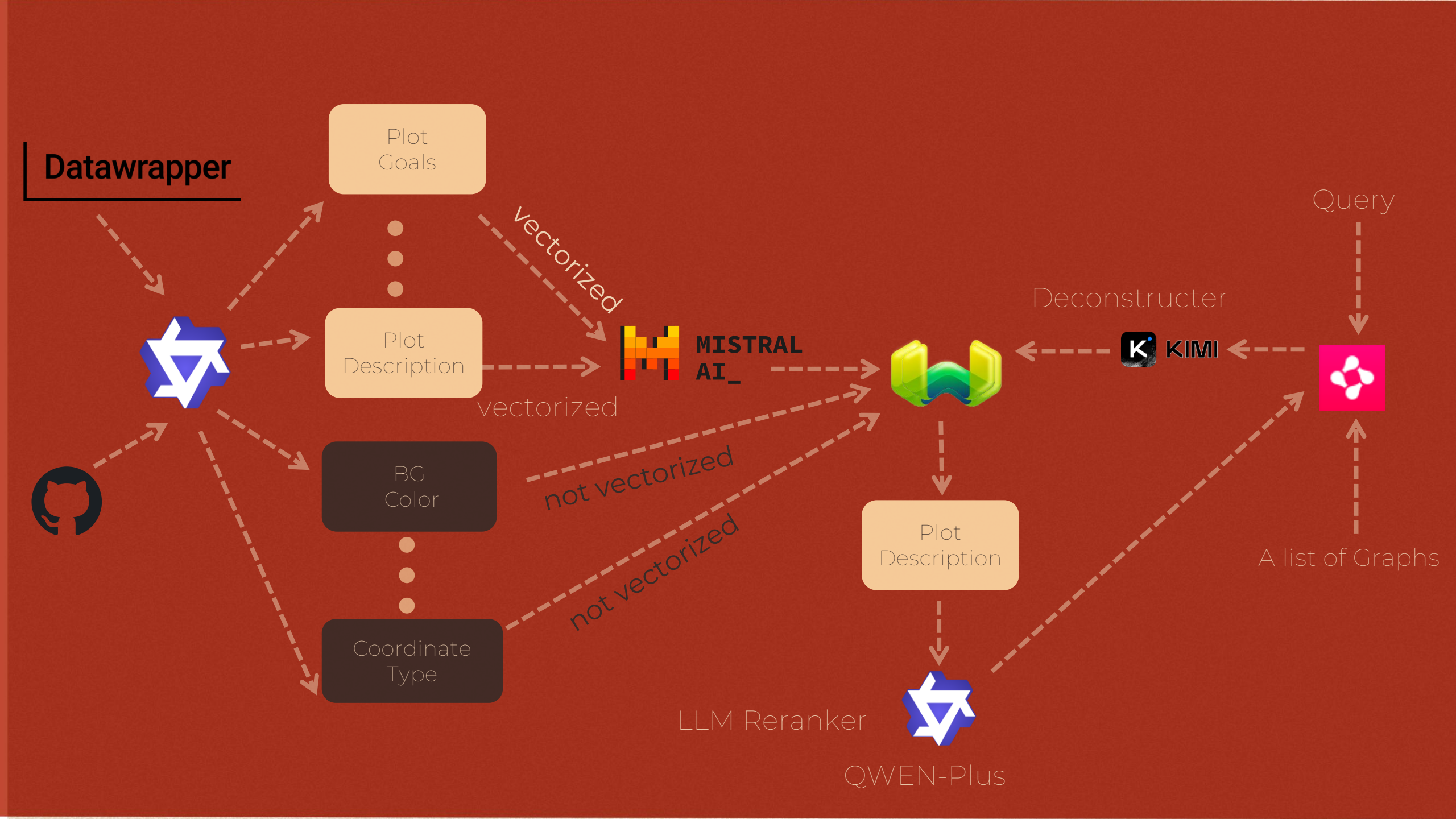



Datawrapper
Plot Goals
vectorized
Plot Description
vectorized
MISTRAL AI_
BG Color
not vectorized
Coordinate Type
not vectorized

prompt_even_more_structured_final-v2-final-12-08-2025.txt



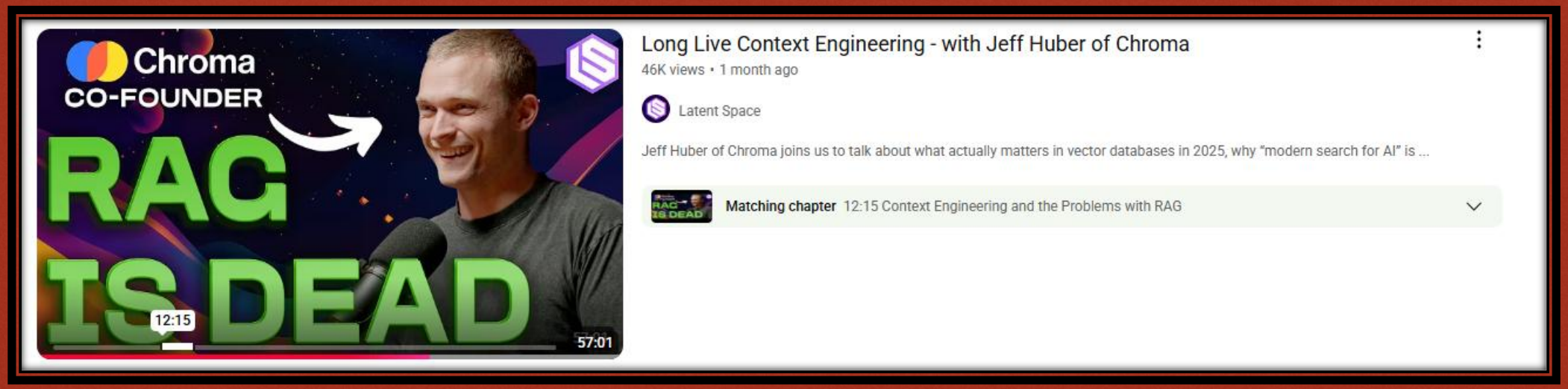
```
**Plot Type**
Primary_Category → (Use terms from the 30DayChartChallenge list where appropriate)
Subcategory → (Use terms from the 30DayChartChallenge list where appropriate)
Specific_Variant_Technique → (Free text)

**Theme & Grid**
Background_Color → (Simple color name)
Background_Type → `light` * `dark`
Grid_Color → (Simple color name)
Grid_Orientation → (MUST use one from: `horizontal`, `vertical`, `both`, `radial`, `none`)
Grid_Layout → (MUST use one from: `single`, `grid`, `faceted`, `small multiples`, `stacked`, `side by side`, `circular`, `radial`, `overlay`, `irregular`, `matrix`)
Grid_Type → (MUST use one from: `major`, `minor`, `implicit`, `subtle`, `reference lines`, `none`)
Grid_Style → (MUST use one or more from: `solid`, `dashed`, `dotted`, `thin`, `light`, `subtle`, `faint`, `minimal`, `none`)

**Legend, Arrangement, Coordinates**
Legend_Position → (MUST use one from: `top left`, `top center`, `top right`, `bottom left`, `bottom center`, `bottom right`, `left`, `right`, `inline`, `embedded`, `inside`, `none`)
Legend_Orientation → (MUST use one from: `horizontal`, `vertical`, `inline`, `embedded`)
Arrangement_Type → (MUST use one or more from: `single`, `grid`, `facets`, `small multiples`, `stacked`, `side by side`, `overlay`, `hierarchical`, `grouped`, `clustered`)
Coordinate_Type → (MUST use one from: `cartesian`, `cartesian_3d`, `polar`, `circular`, `geographic_general`, `flow_network`, `schematic`, `linear`, `categorical`, `logarithmic`, `mixed`, `none`)

**Typography**
Title_Font_Family → (Font name only)
Title_Font_Type → (MUST use one from: `sans-serif`, `serif`, `slab serif`, `monospace`, `script`, `handwritten`, `blackletter`)
Title_Style → (MUST use one or more from: `regular`, `bold`, `italic`, `semi-bold`, `medium`, `extra bold`, `heavy`, `mixed`)
Subtitle_Font_Family → (Font name only)
Subtitle_Font_Type → (MUST use one from the list above)
Subtitle_Style → (MUST use one or more from the list above)
Axis_Labels_Font_Family → (Font name only)
Axis_Labels_Font_Type → (MUST use one from the list above)
Axis_Labels_Style → (MUST use one or more from the list above)

**Color**
Palette_Type → (MUST use one or more from: `sequential`, `diverging`, `qualitative`, `categorical`, `monochrome`, `grayscale`, `semantic`, `brand colors`, `highlight`, `accent`, `mixed palette`)
Number_of_Distinct_Colors → (Integer only)

**Data Source**
Source_Name → (Free text)
Year → (Free text)

**Assessments**
Readability_Assessment → (MUST use one from: `very_high`, `high`, `medium_high`, `medium`, `medium_low`, `low`, `mixed`, `none`)
Readability_Assessment_Explanation → (One-sentence justification)
Color_Accessibility_Assessment → (MUST use one from: `excellent`, `very_good`, `good`, `moderate`, `fair`, `poor`, `not_applicable`, `none`)
Color_Accessibility_Assessment_Explanation → (One-sentence justification)
Text_Contrast_Ratio_WCAG → (MUST use one from: `aaa_compliant`, `aa_compliant`, `aa_borderline`, `below_aa`, `not_assessed`, `not_applicable`, `none`)
Text_Contrast_Ratio_WCAG_Explanation → (One-sentence justification)

%%%

Section_1_Data_and_Variable_Types@
Miles_per_Gallon: Continuous
Horsepower: Continuous
Weight: Continuous
Origin: Categorical
Model_Year: Temporal

%%%

Section_2_Variable_Mapping@
Horsepower is mapped to the x-axis via `aes(x = horsepower)`. Miles_per_Gallon is mapped to the y-axis via `aes(y = mpg)`. Origin is mapped to both color via `aes(color = origin)` and facet panels
(~origin)`.

%%%

Section_3_Color_Encoding_Details@
Key_Colors → Teal #448154*Green #21918c*Yellow #fde725
Semantic_Meaning → Color maps to country of origin using the `scale_color_viridis_d()` function.

%%%

Section_4_Chart_Element_Identification@
Statistical_Summaries → trend_lines*correlation_patterns

%%%

Section_5_Layout_Details@
Structure_Description → A 1x3 grid of scatter plots, arranged horizontally as defined by `facet_wrap` with `ncol=3`.
Aspect_Ratio → Each facet is approximately square.

%%%

Section_6_Axes_and_Scales@
X-Axis_Horsepower:
Scale_Type → Linear
Position → Bottom
Y-Axis_Miles_per_Gallon:
Scale_Type → Linear
Position → Left

%%%

Section_7_Annotation_and_Storytelling_Elements@
Annotations → trend_annotations*contextual_info
```

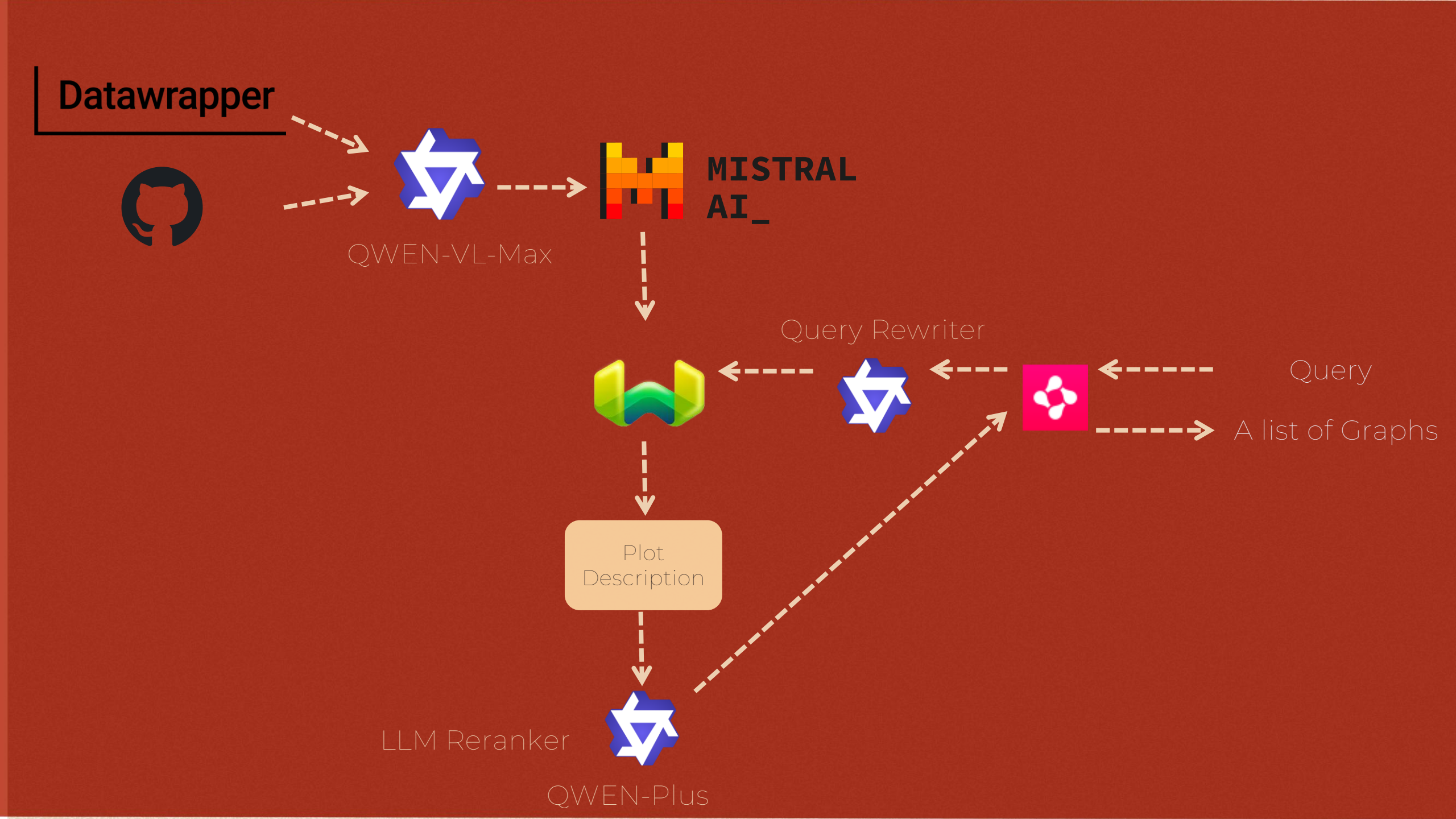

Datawrapper
Plot Goals
vectorized
Plot Description
MISTRAL AI_
vectorized
BG Color
not vectorized
Coordinate Type
not vectorized
Deconstructer
KIMI
Query
Plot Description
A list of Graphs
LLM Reranker
QWEN-Plus

MODE 3

USING LONGER CONTEXT

RAG IS DEAD!

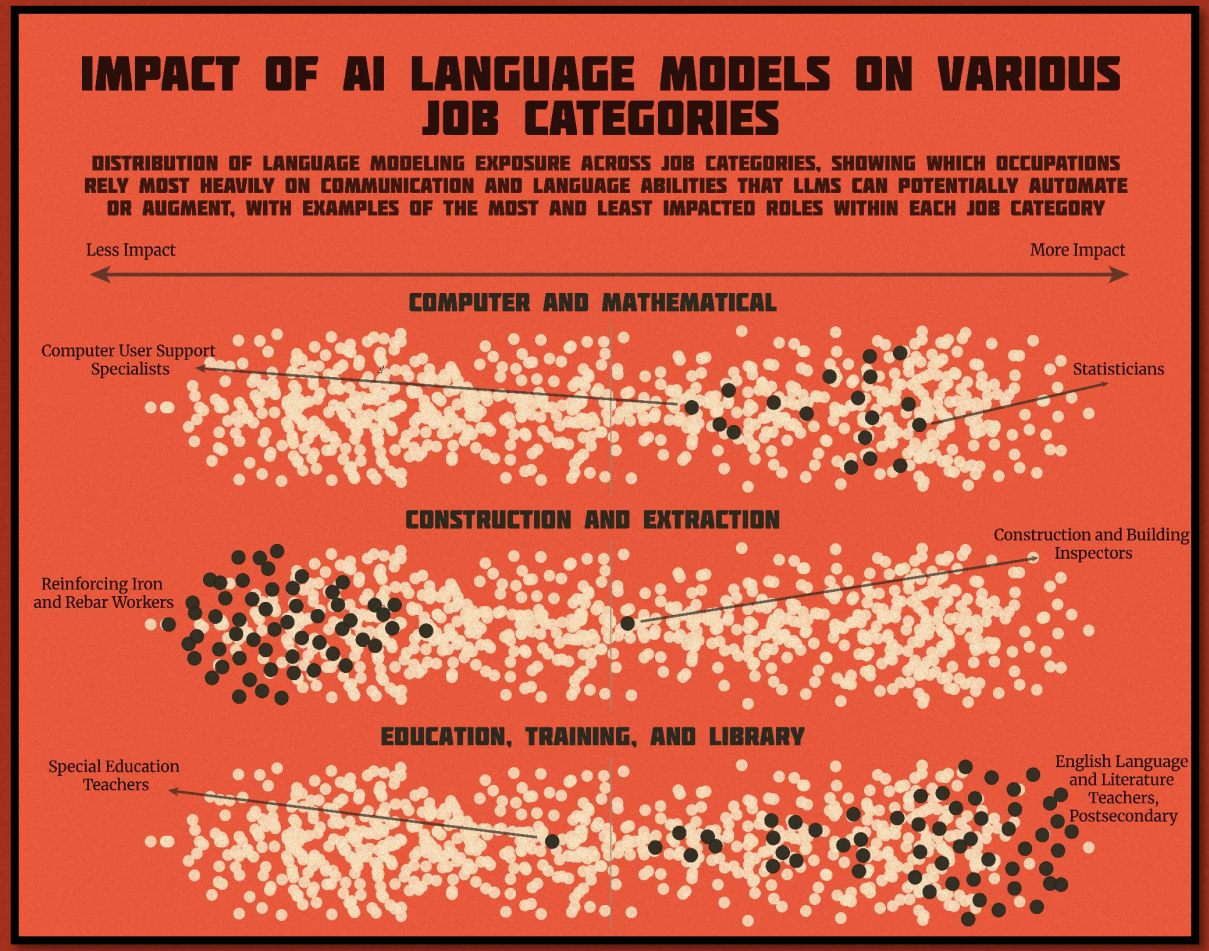

Chroma
CO-FOUNDER
RAG
IS DEAD
12:15
57:01
Long Live Context Engineering - with Jeff Huber of Chroma
46K views • 1 month ago
Latent Space
Jeff Huber of Chroma joins us to talk about what actually matters in vector databases in 2025, why "modern search for AI" is ...
Matching chapter 12:15 Context Engineering and the Problems with RAG


Datawrapper
QWEN-VL-Max
MISTRAL AI_
Query Rewriter
Query
A list of Graphs
Plot Description
LLM Reranker
QWEN-Plus

LIMITATIONS

Garbage In, Garbage Out

The Challenge of Subjectivity



Knowing What You Didn't Find

Evaluation is Hard

IMPROVEMENTS

Better prompt, better VLM model

Better toolcalling logic

Better query rewriting

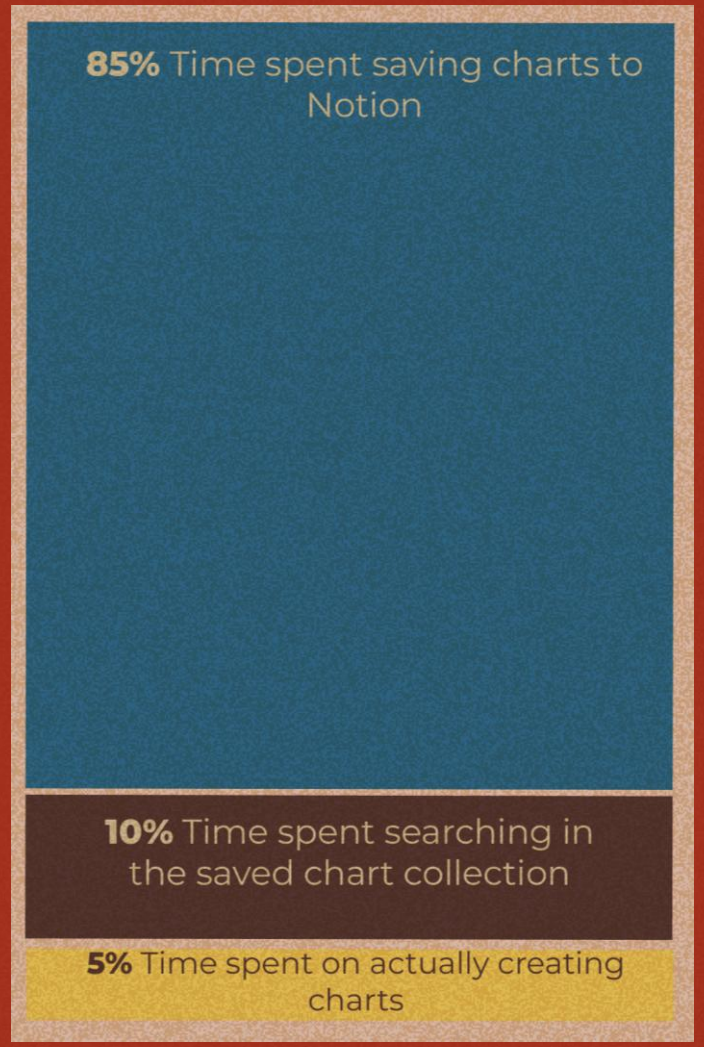
TAKEAWAY

TAKEAWAY

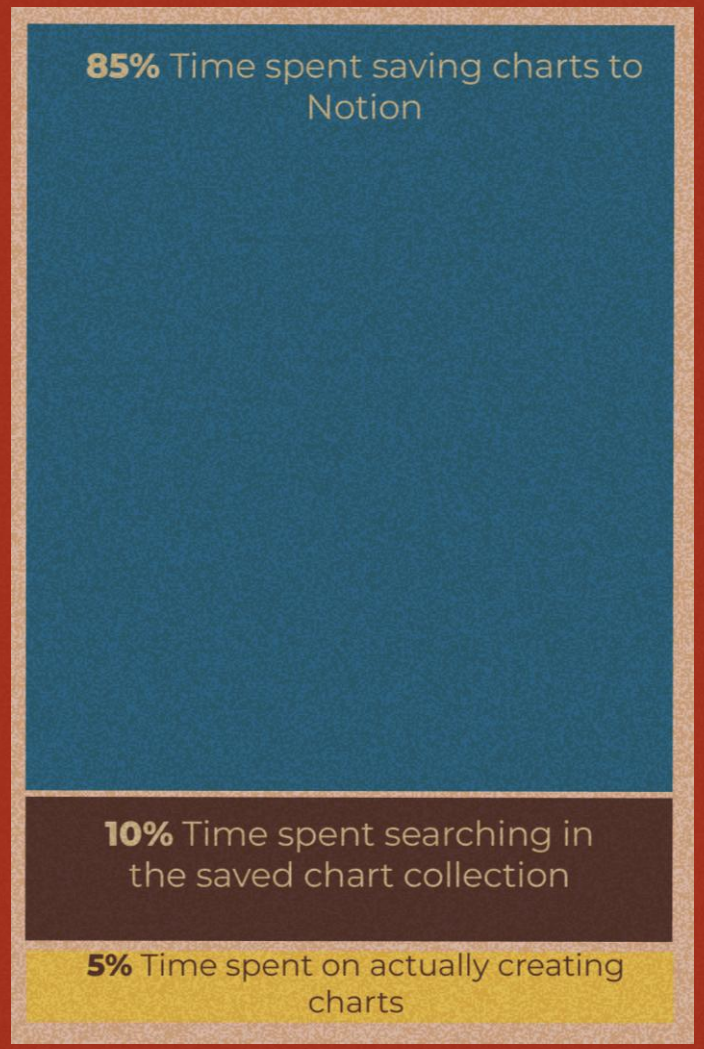
“Stop using your head as an open browser tab, instead use RAGs”

WHAT IS NEXT ...

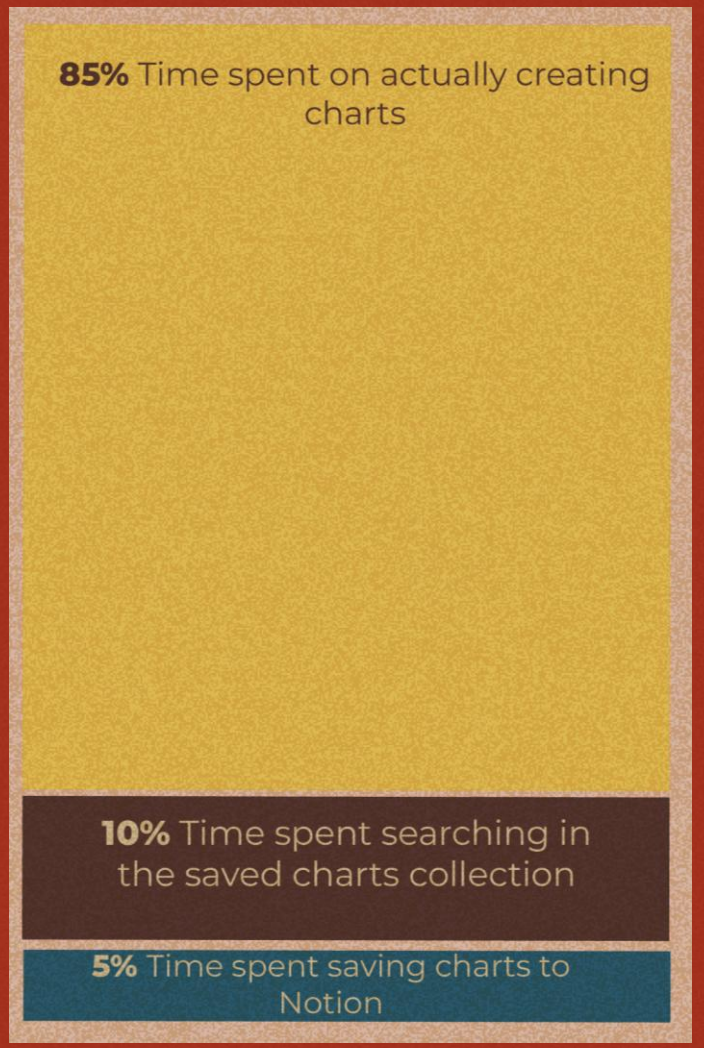
How it started



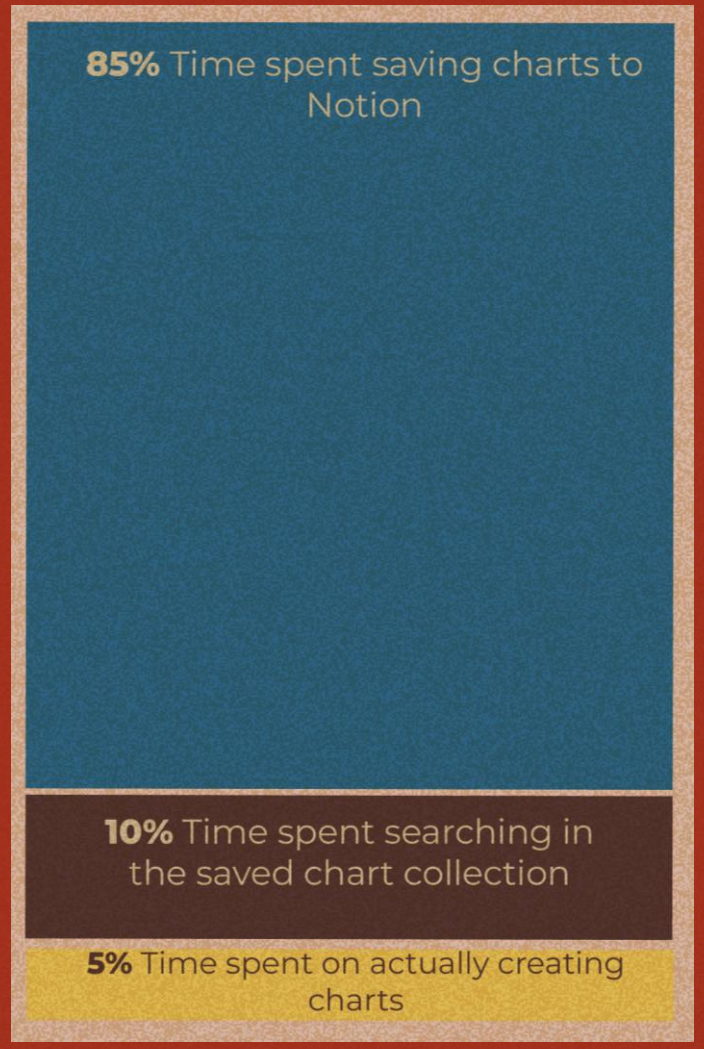
How it started



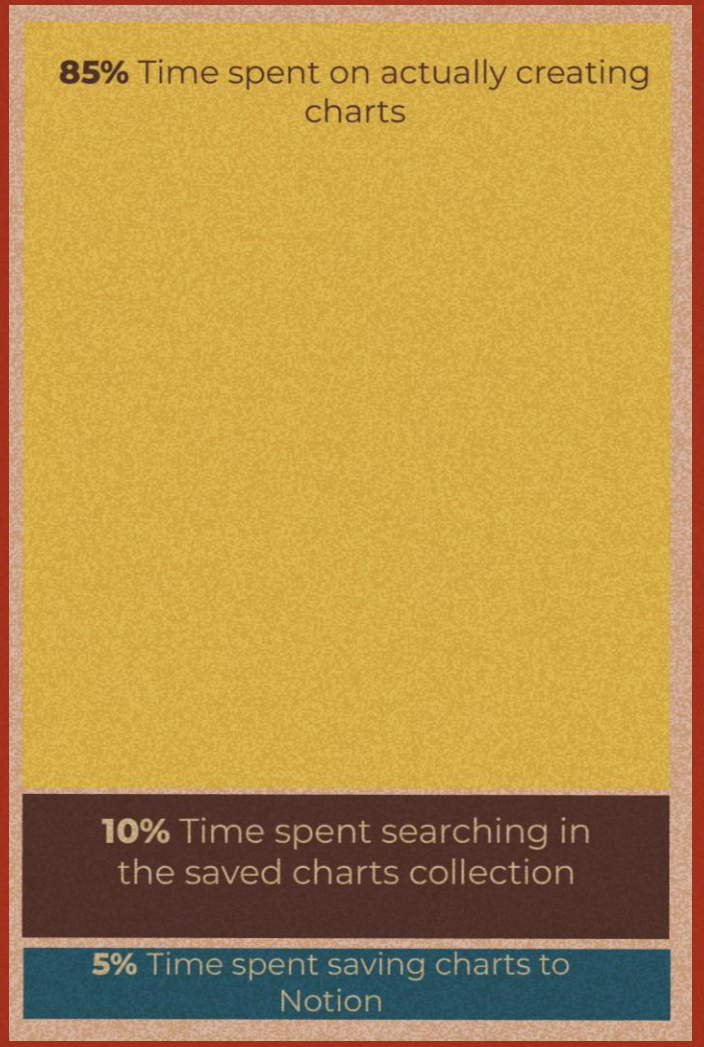
How I expected it to go



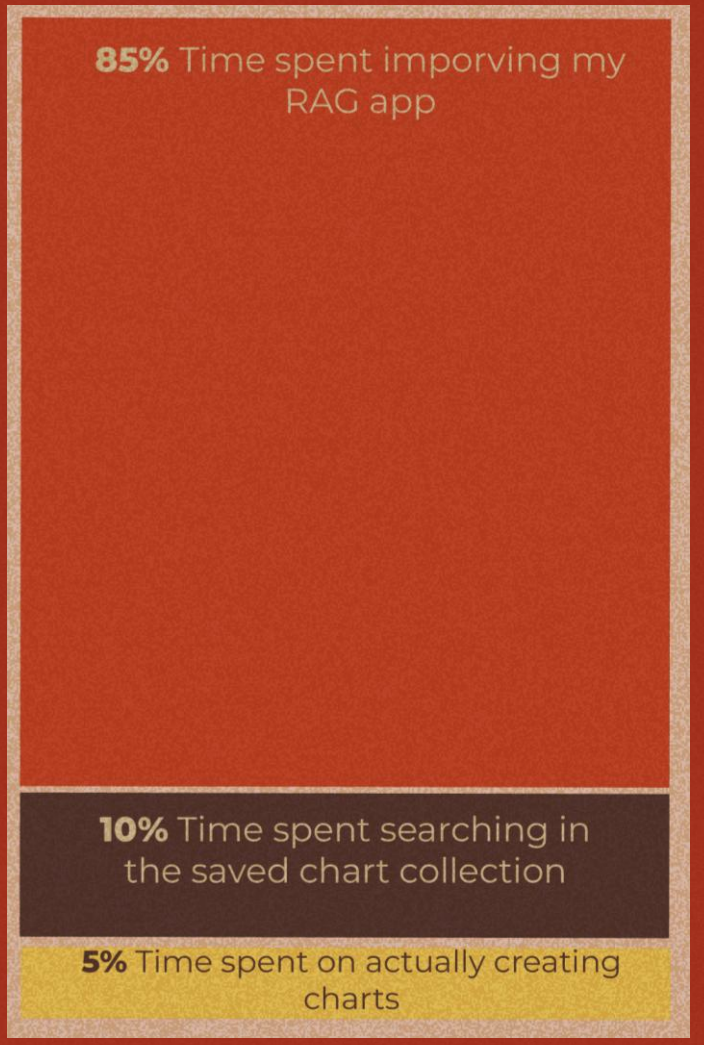
How it started



How I expected it to go



How it's going



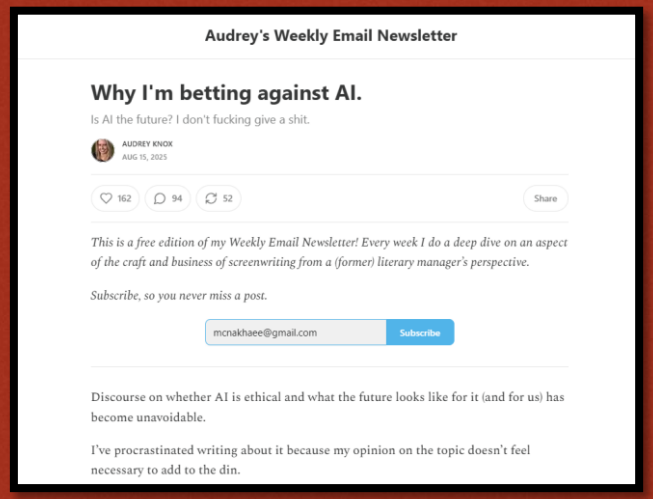
THANK YOU

APPENDIX

Using AI in Creative work

Never AI camp

Vibe Artists



Audrey's Weekly Email Newsletter

Why I'm betting against AI.

Is AI the future? I don't fucking give a shit.

AUDREY KNOX
AUG 15, 2025

162 94 52 Share

This is a free edition of my Weekly Email Newsletter! Every week I do a deep dive on an aspect of the craft and business of screenwriting from a (former) literary manager's perspective.

Subscribe, so you never miss a post.

mcnakhaee@gmail.com Subscribe

Discourse on whether AI is ethical and what the future looks like for it (and for us) has become unavoidable.

I've procrastinated writing about it because my opinion on the topic doesn't feel necessary to add to the din.

[Why I'm betting against AI.]

The success of AI music creators sparks a debate on the future of the music industry

People complain "that you're using a computer to do all the work for you. I don't see it that way. I see it as any other tool that we have," said Scott Smith, whose AI band, Pulse Empire, was inspired by 1980s British synthesizer-driven groups like New Order and Depeche Mode.

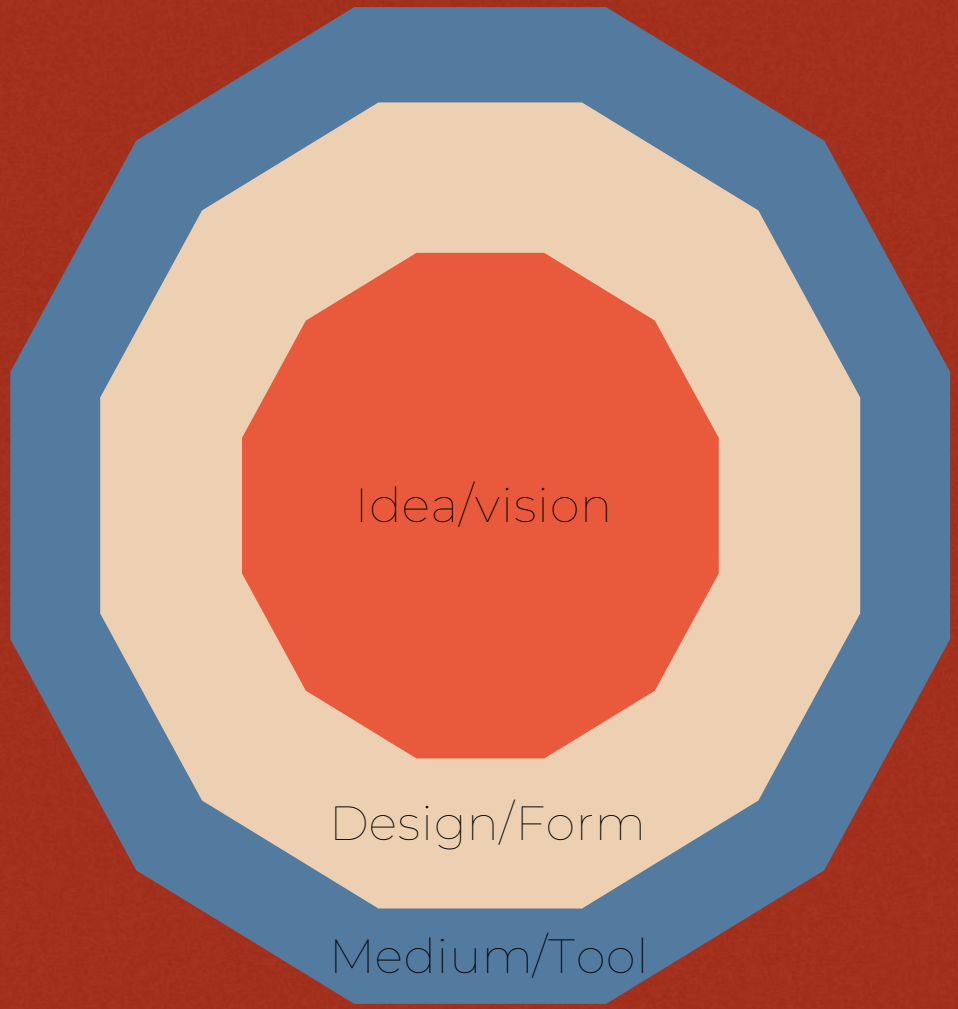
Smith, 56 and a semi-retired former U.S. Navy public affairs officer in Portland, Oregon, said "music producers have lots of tools in their arsenal" to enhance recordings that listeners aren't aware of.

Like McCann, Smith never mastered a musical instrument. Both say they put lots of time and effort into crafting their music.

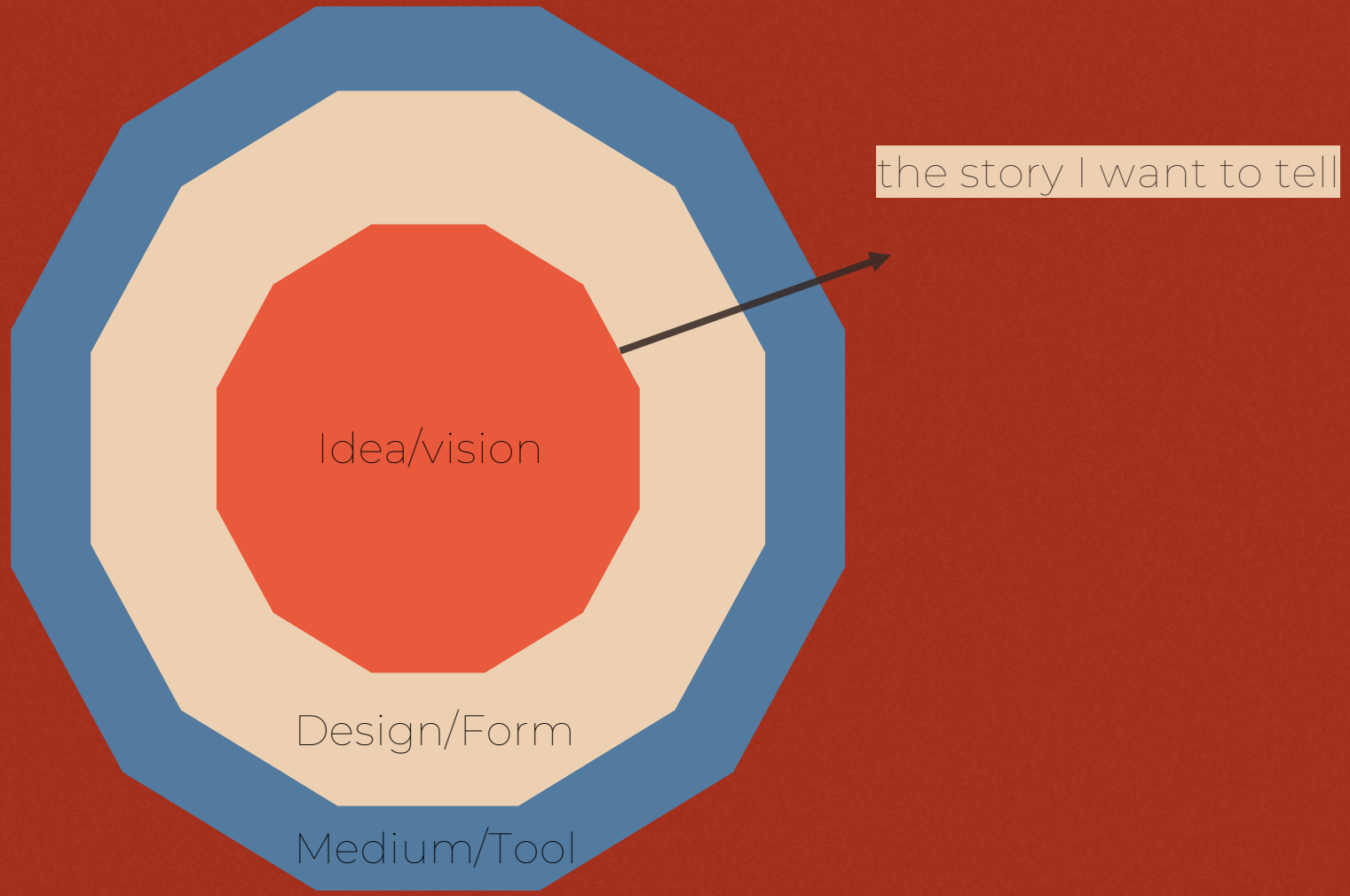
Once Smith gets inspiration, it takes him just 10 minutes to write the lyrics. But then he'll spend as much as eight to nine hours generating different versions until the song "matches my vision."

[The success of AI music creators sparks a debate on the future of the music industry]

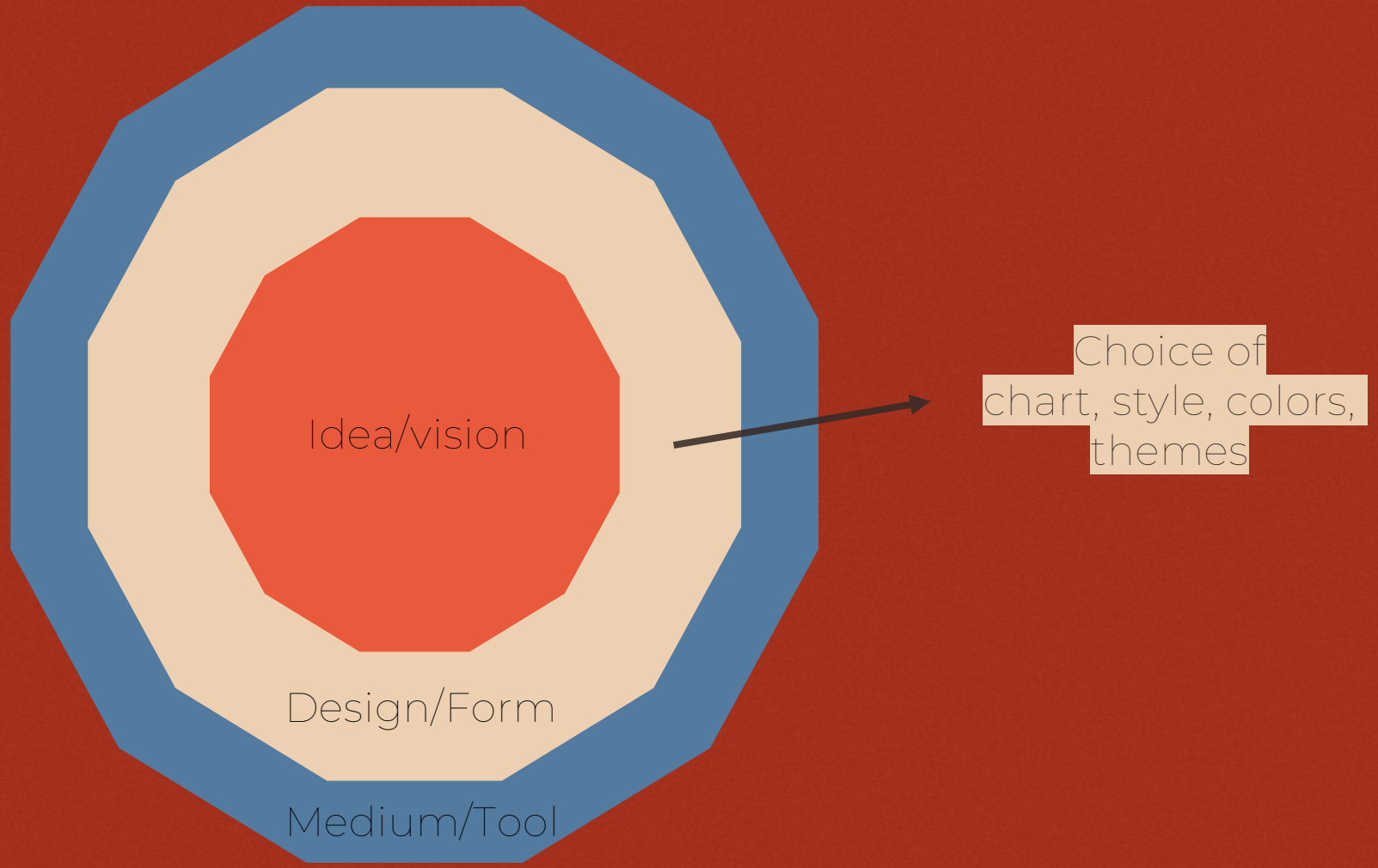
Creativity Blob



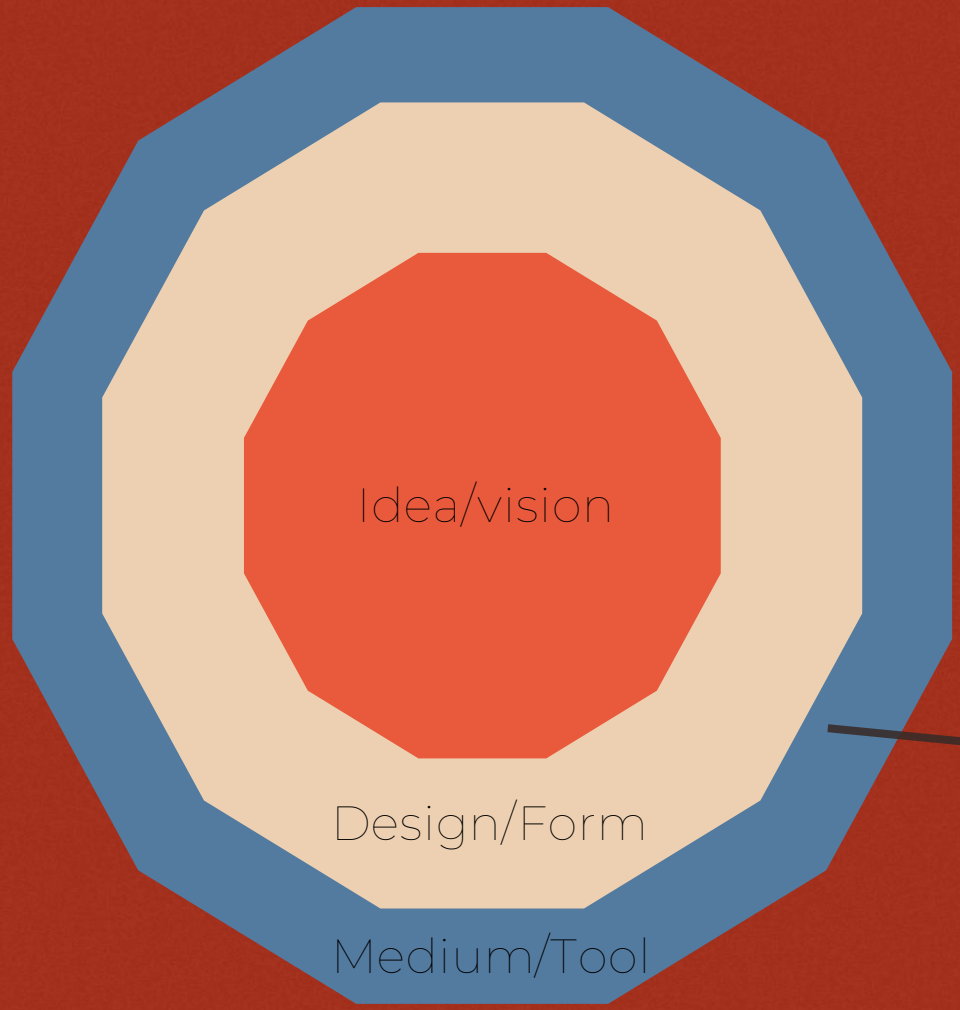
Creativity Blob



Creativity Blob

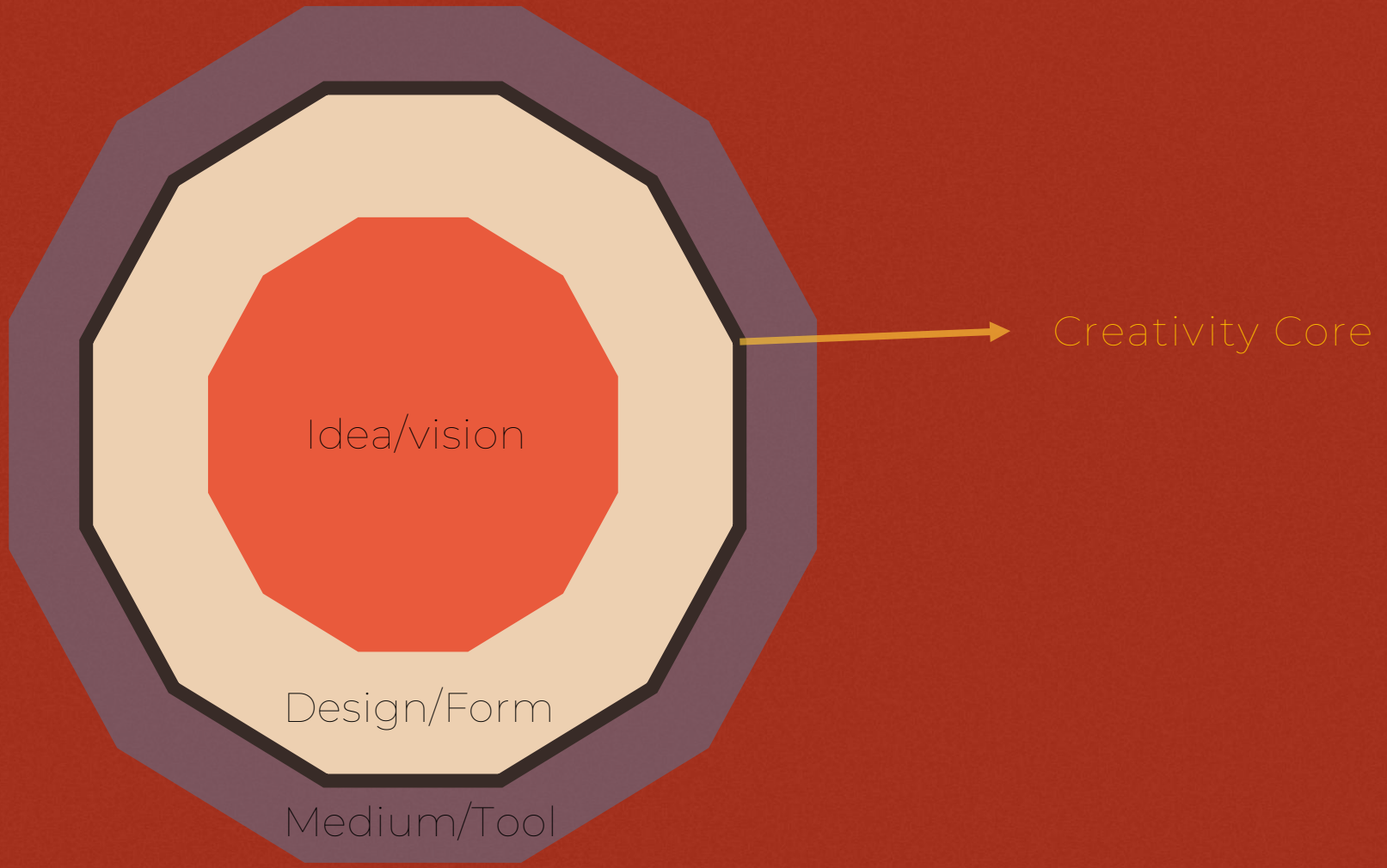


My Rule

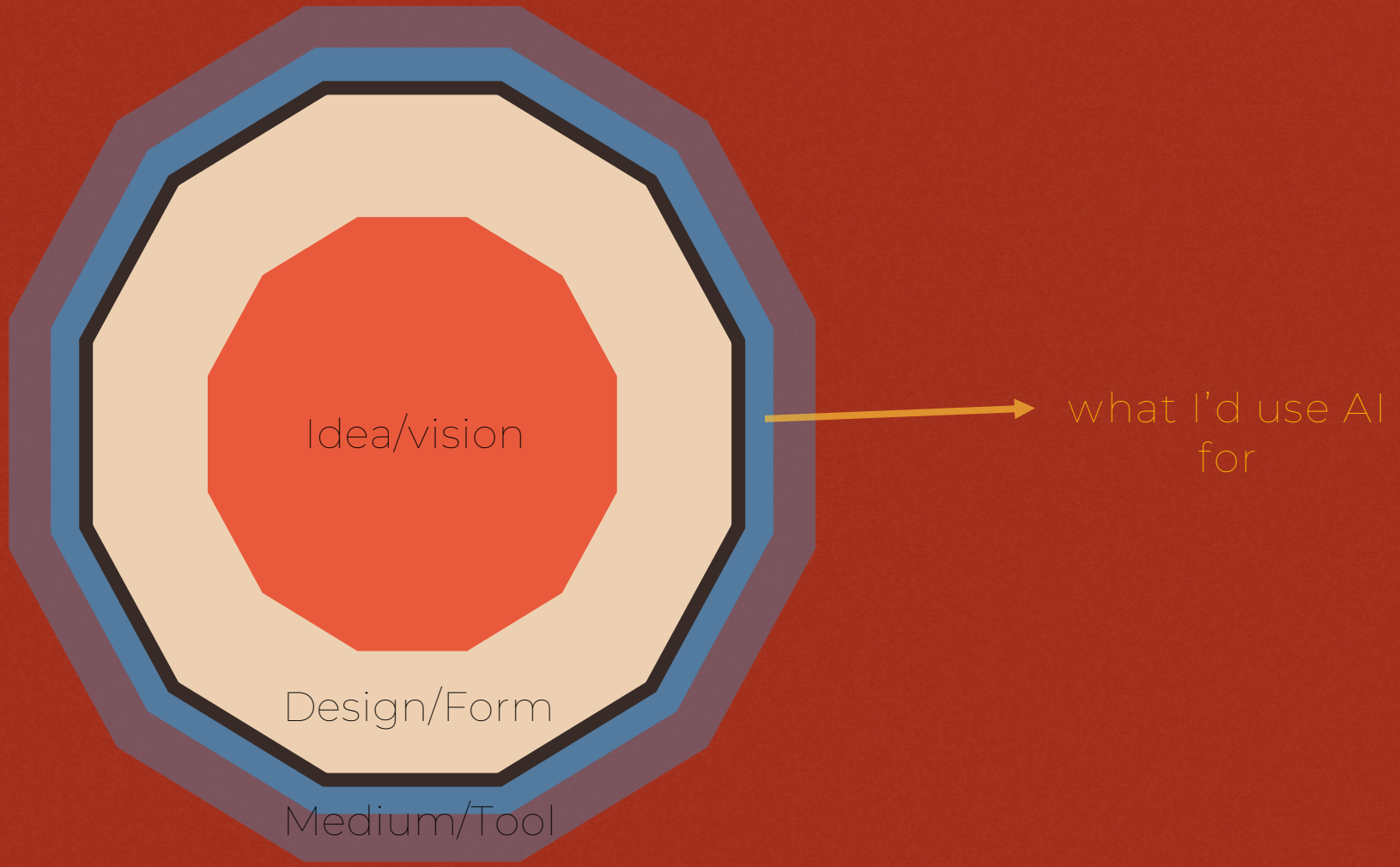


ggplot2, Illustrator, plotnine, etc.

Creativity Blob



Creativity Blob



My Rule-of Thumb for using AI

Where I'd use AI ☑

For my data research and exploring data sources

To write the entire data scraping code

To proofread my title, subtitles, captions, etc.

Where I draw a line in using AI ✕

To give me inspirations, ideas or brainstorm with me

To decide which variables to visualize

To recommend me chart types.

To rocommend or pick themes, colors, and overall aesthetics for me

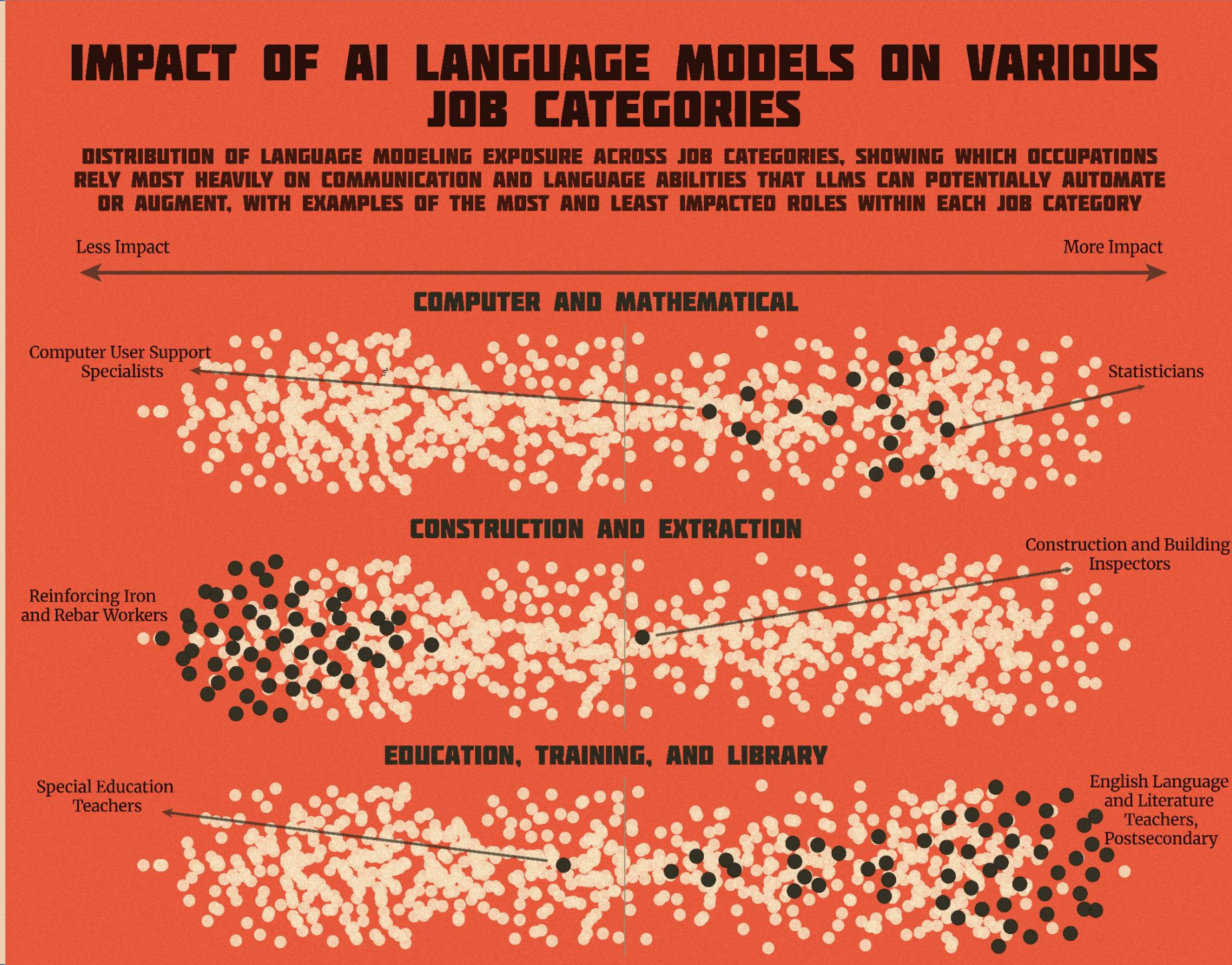
To generate the entire ggplot2 code

Data

Each raw/point = occupation

Variables:

- Job Category
- Impact Score
- Job Title

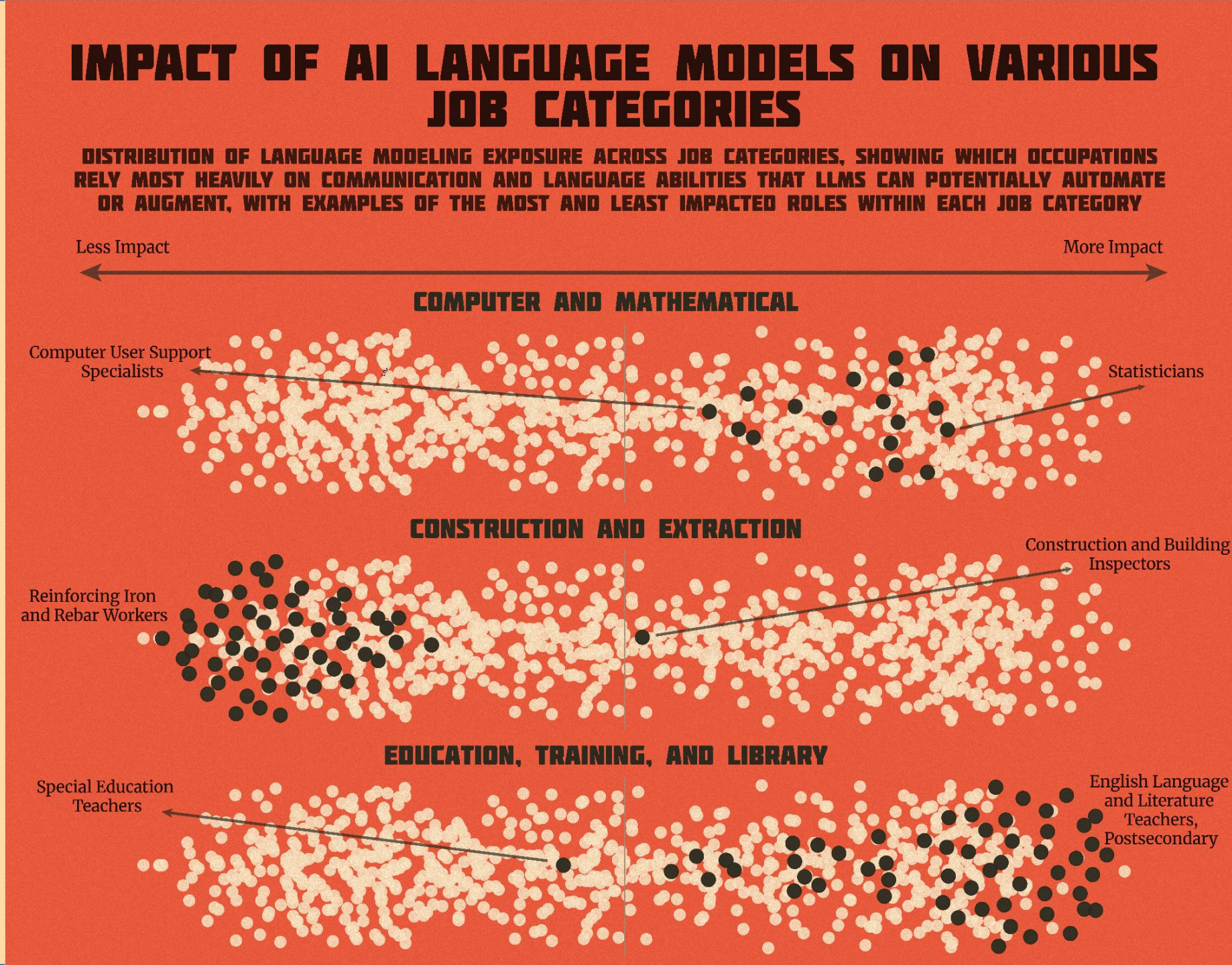


Aesthetics

x-axis → Impact score

y-axis → random position

color/fill → Two groups: highlighted jobs (black) vs. general jobs (white).

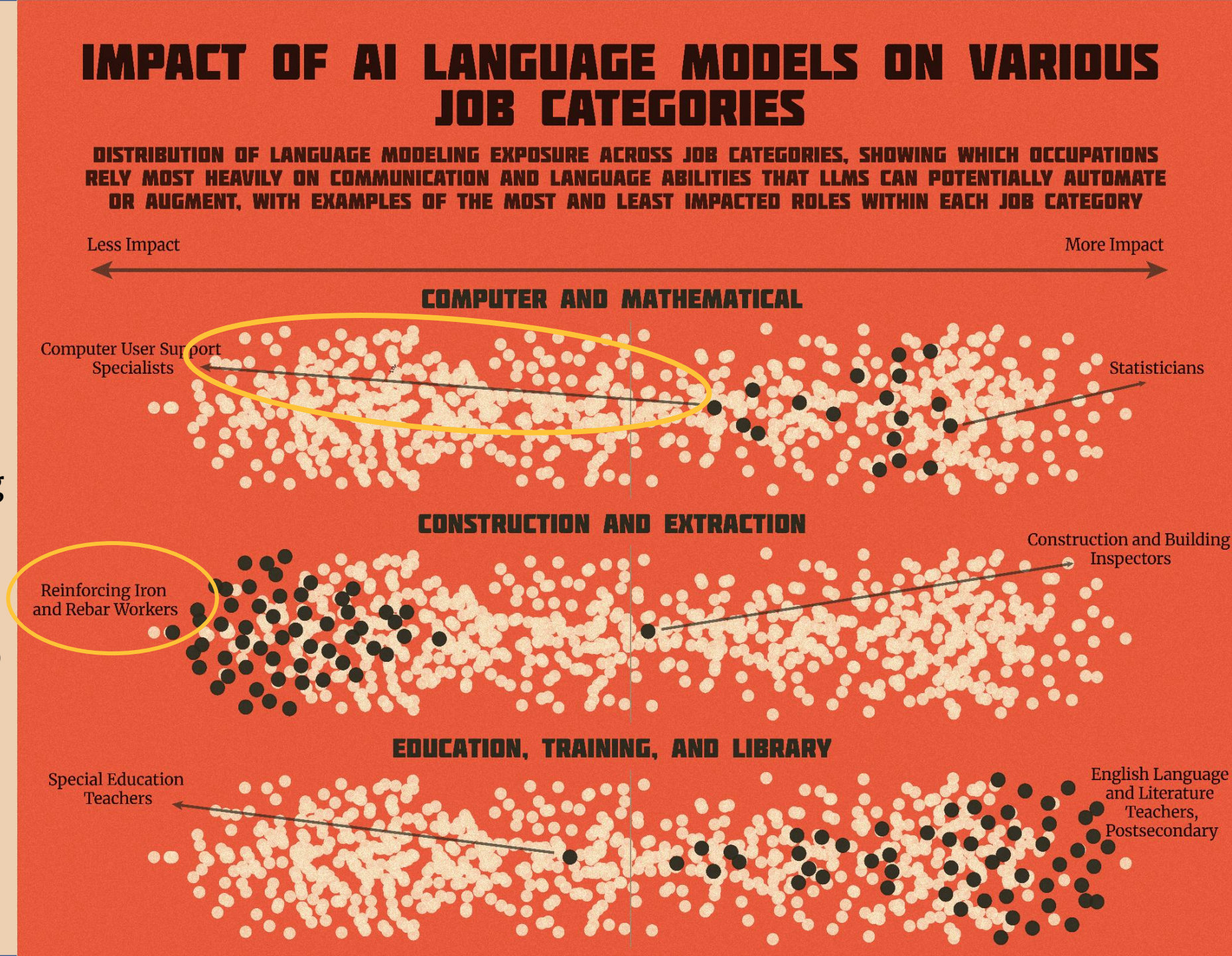


Geometries

geom_point() → For scatter of jobs within categories.

geom_text() → For labeling specific jobs.

geom_segment() → For arrows connecting labels to points.



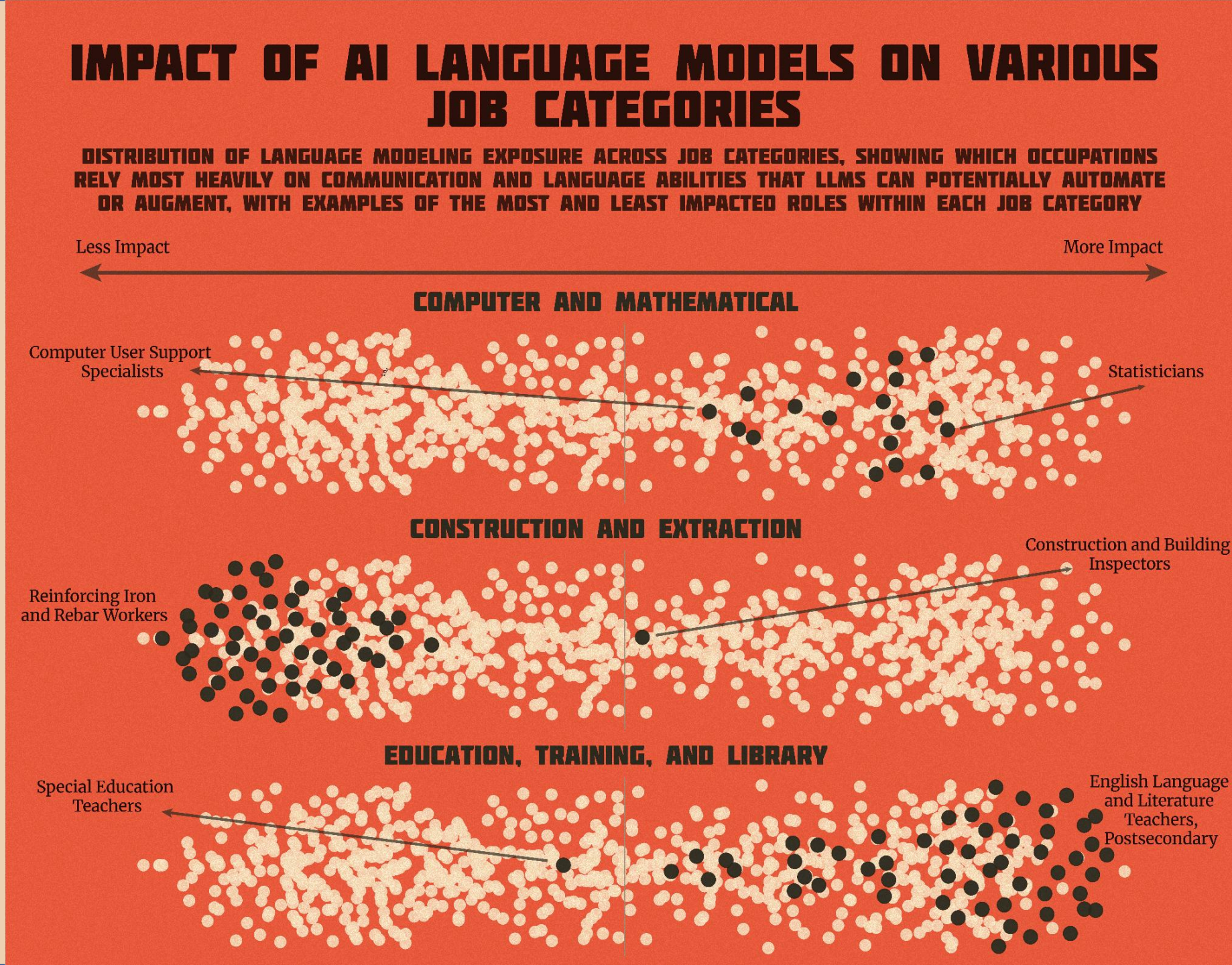
Scales

x scale: Continuous

color scale: Categorical

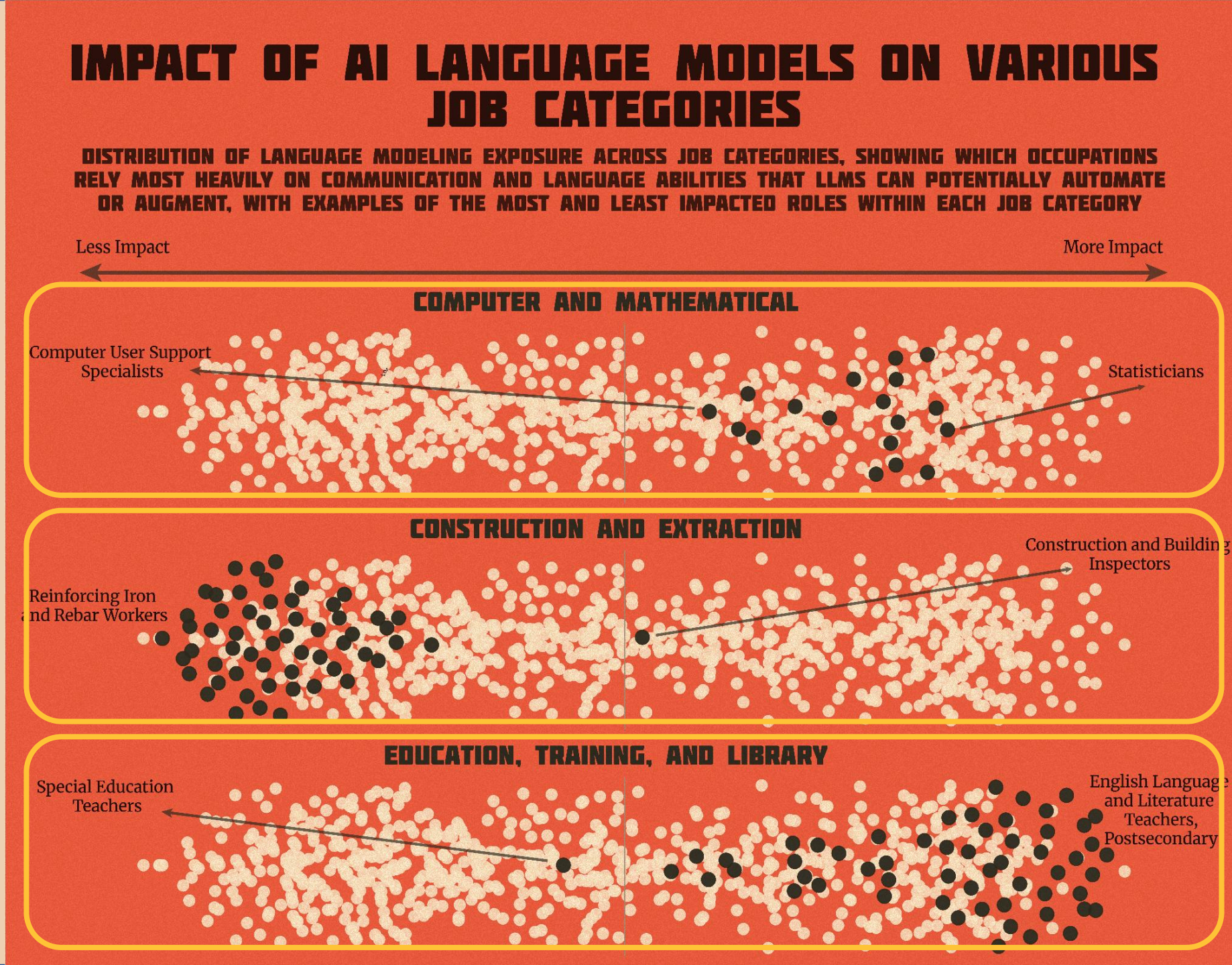
(highlighted = black,

non-highlighted = white)



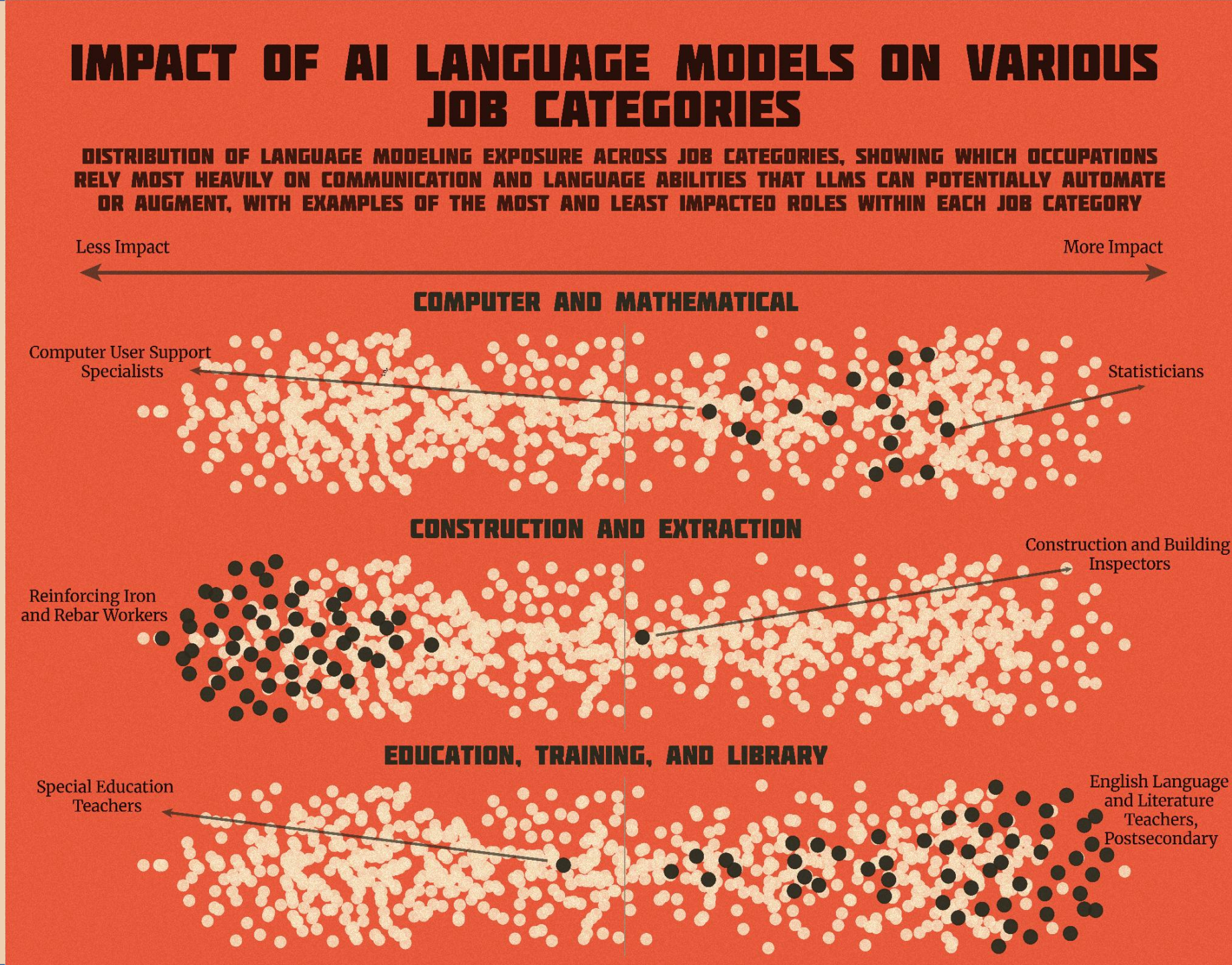
Facets

3 major facets/panels



Coordinate

coord_cartesian(): standard Cartesian coordinates



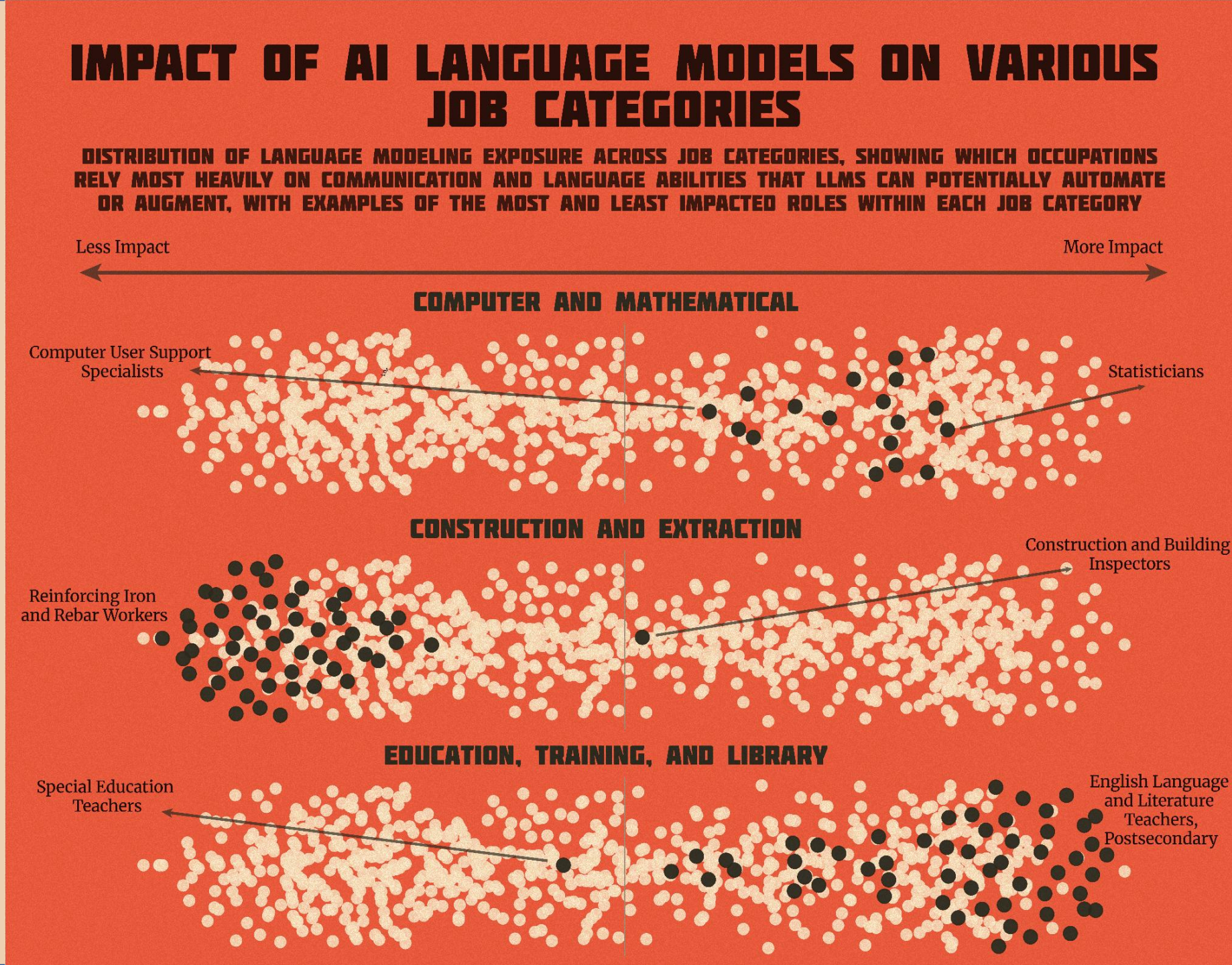
Theme

Background: #E95A3C

Font family (**MOLOT**, Merriweather)

Font color (black),

Font style (**bold**)



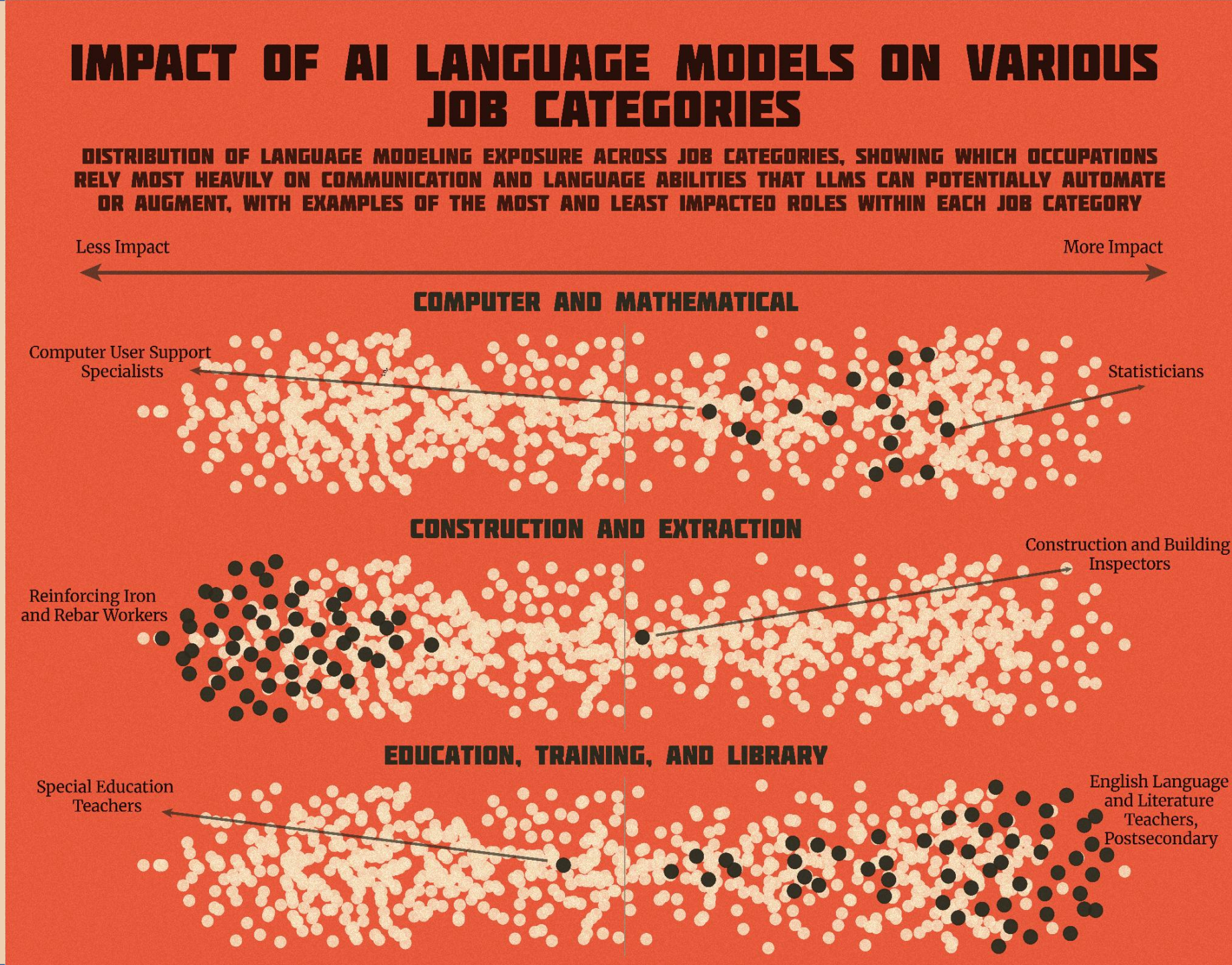
Scales

x scale: Continuous

color scale: Categorical

(highlighted = black,

non-highlighted = white)



Aesthetics

x-axis → Impact score

y-axis → random position generated by the beeswarm algorithm

color/fill → Two groups: highlighted jobs (black) vs. general jobs (white).