Alexander Lex
http://vdl.sci.utah.edu

THE UNIVERSITY OF UTAH

The reVISit User Study Platform and Applications in Studying Misinformation

visualization design lab

SCI
www.sci.utah.edu

visualization
design lab

http://vdl.sci.utah.edu/

TECHNICAL CONTRIBUTIONS

Novel Visualization Techniques

Visualization Process Innovations

Data Wrangling Methods

Accessibility

DOMAIN DRIVEN TECHNIQUES

Tailored Methods and Systems for High Impact Science Problems

EMPIRICAL & THEORETICAL WORK

Evaluation Methodology

Interview Studies

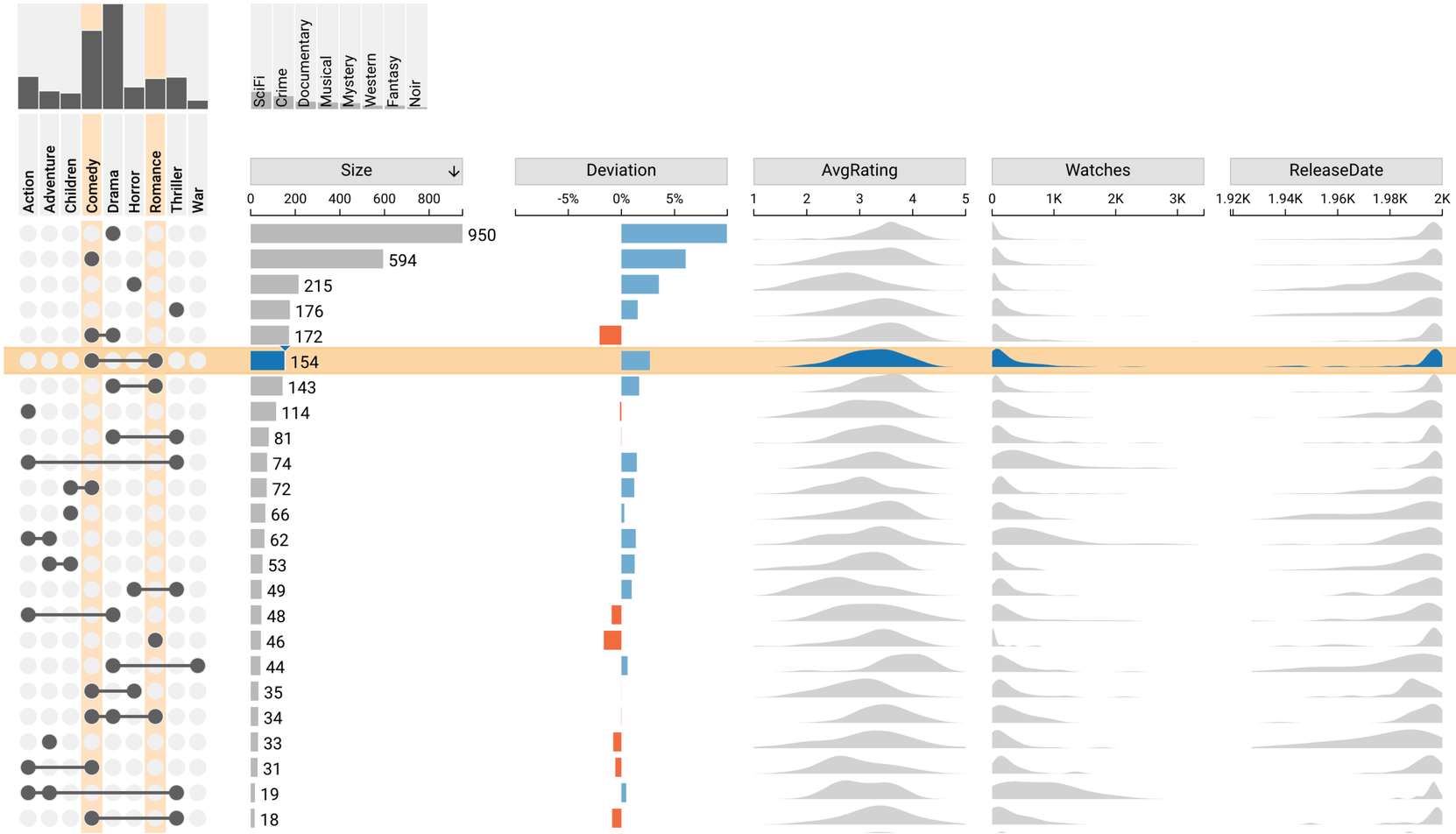
Visual Misinformation

Accessibility

TECHNICAL CONTRIBUTIONS

Set Data - UpSet

Novel Visualization Techniques

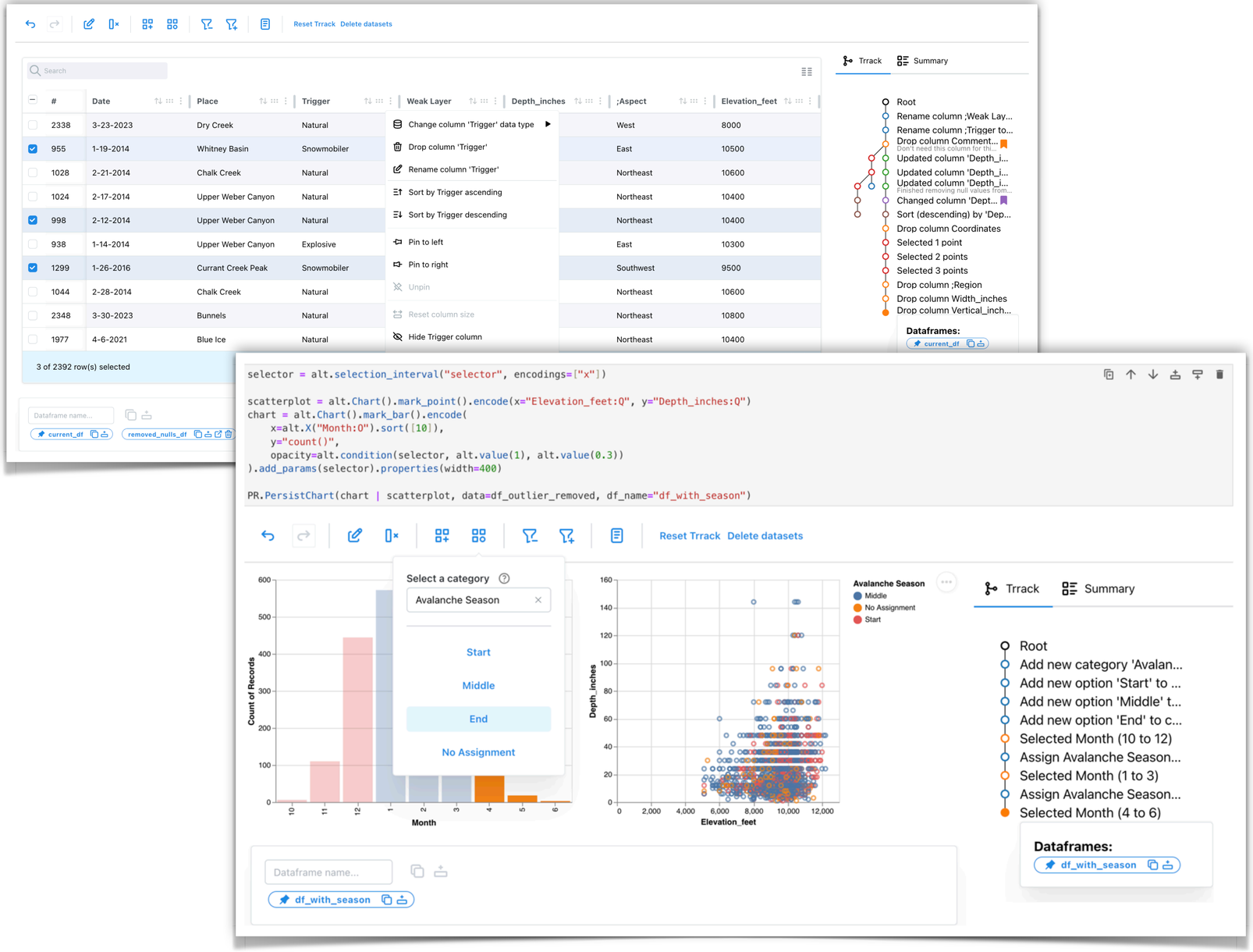


TECHNICAL CONTRIBUTIONS

Vis in Notebooks & Provenance

Visualization Process Innovations

Kiran Gadhave

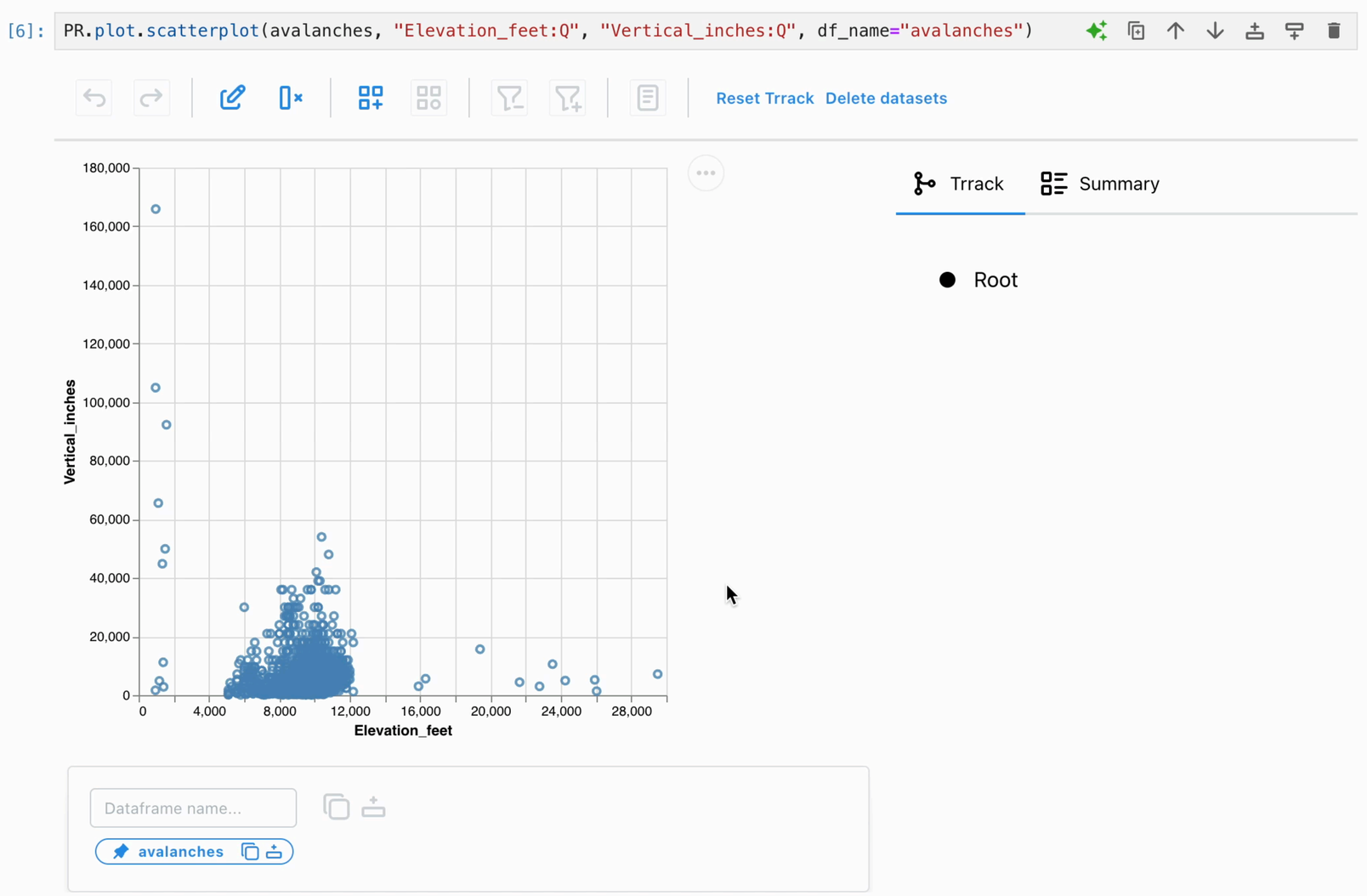


TECHNICAL CONTRIBUTIONS

Vis in Notebooks & Provenance

Visualization Process Innovations

Kiran Gadhave



TECHNICAL CONTRIBUTIONS

Data Wrangling Methods

DOMAIN DRIVEN TECHNIQUES

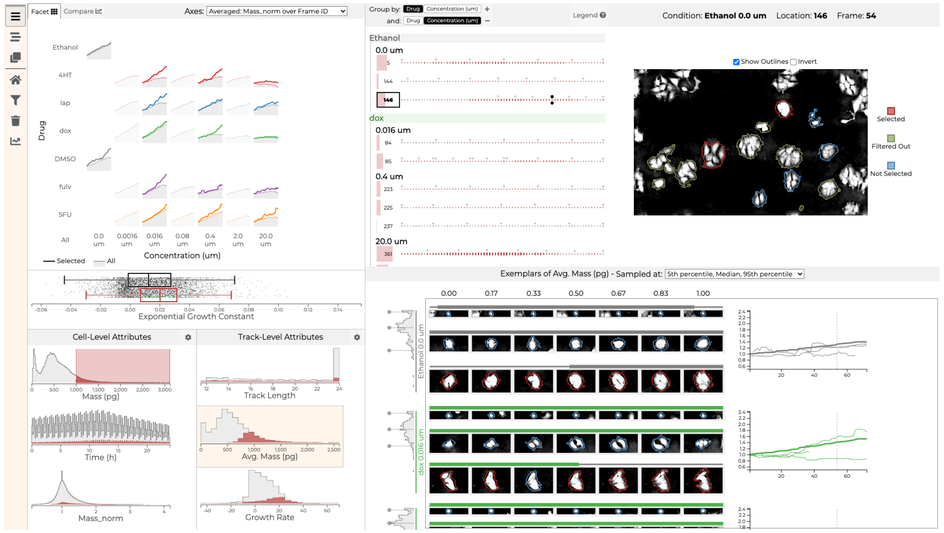
Tailored Methods and Systems for High Impact Science Problems

DOMAIN DRIVEN TECHNIQUES

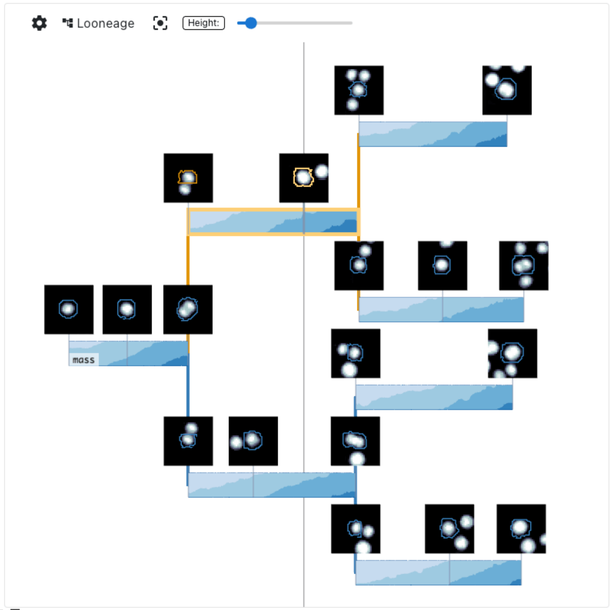
Cell Microscopy

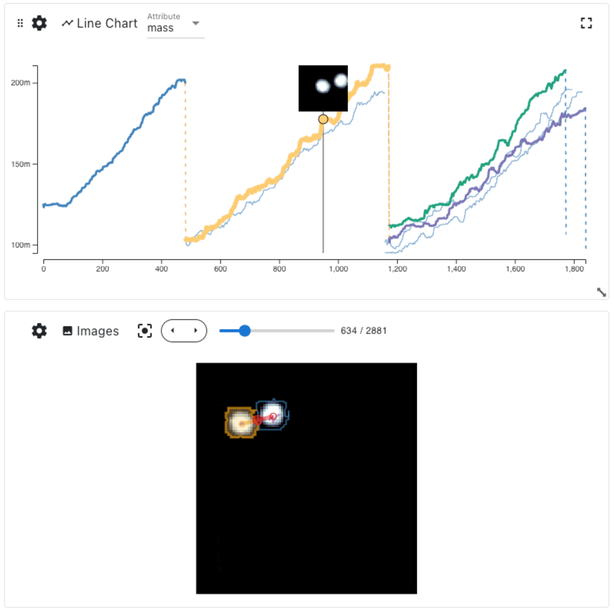
Tailored Methods and Systems for High Impact Problems

Devin Lange



[Honorable Mention at VIS 21]





[Best Paper at VIS 25]

DOMAIN DRIVEN TECHNIQUES

Tailored Methods and Systems for High Impact Science Problems

EMPIRICAL & THEORETICAL WORK

Evaluation Methodology

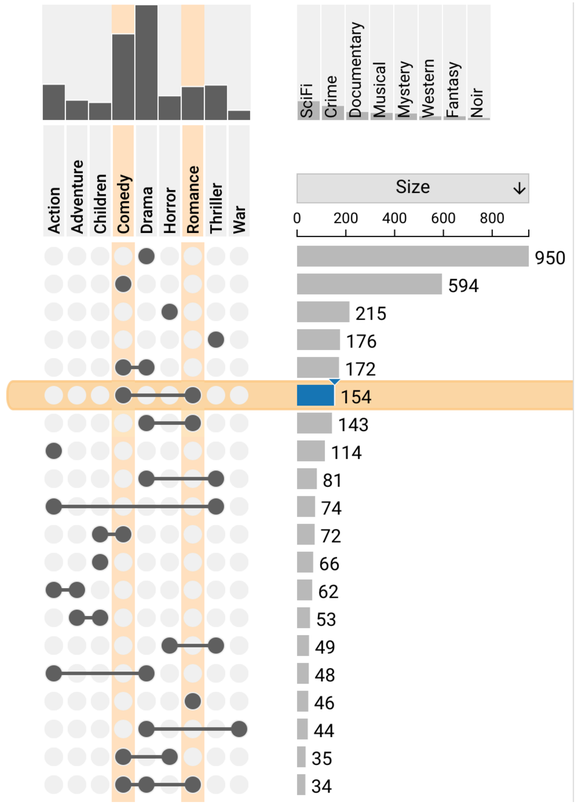
Interview Studies

Visual Misinformation

Accessibility

Text Descriptions & Physical Teaching Aids

Accessibility



Movies are Purists!

Most movies stick to one genre. The four largest intersections are for movies that are in just one genre. There are a lot of Comedy-Drama combinations, but more interestingly, there are also a lot of romantic comedies.

This UpSet plot shows movie genre overlap. The sets are movie genres. The items are movies.

Text Description

UpSet Introduction

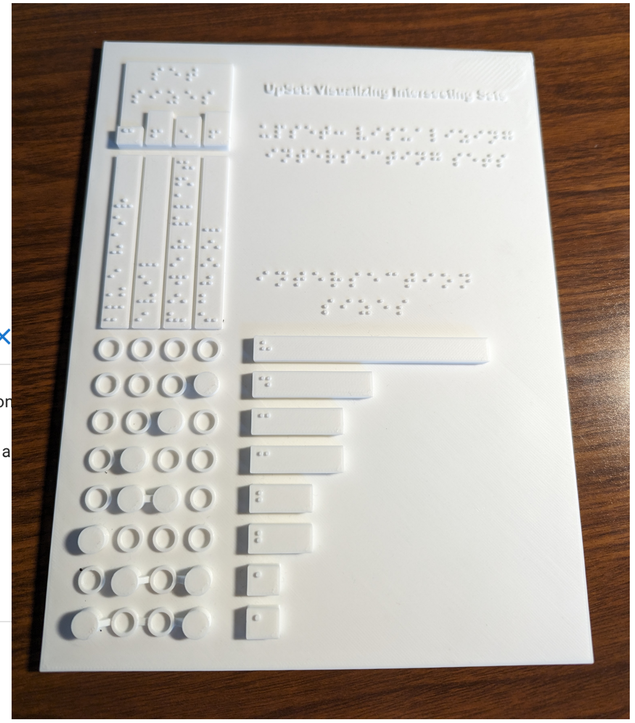
This is an UpSet plot that shows the intersections of 9 sets. To learn about UpSet plots, visit https://upset.app. The largest 2 intersections are just Drama (950) and just Comedy (594). The largest intersection of at least two sets is between Comedy and Drama, with 172 movies. Other large intersections also involve Comedy and Horror.

Dataset Properties

The dataset shows attributes of movie genre overlap. The dataset contains 17 sets and 6303 elements, of which 9 sets are shown in the plot.

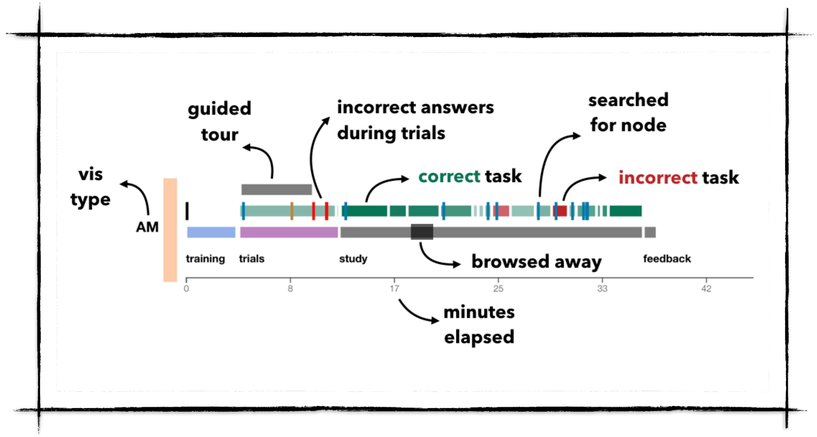
Set Properties

- The set sizes are diverging a lot, ranging from 143 to 1603.

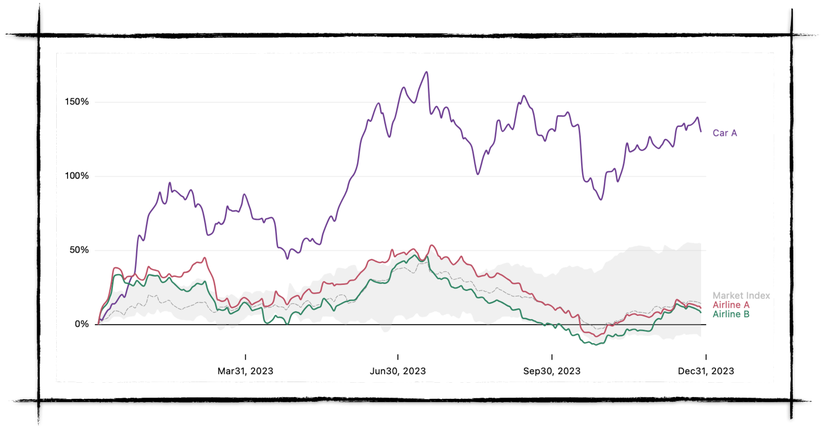


EMPIRICAL & THEORETICAL WORK

Evaluating Interactive Visualizations

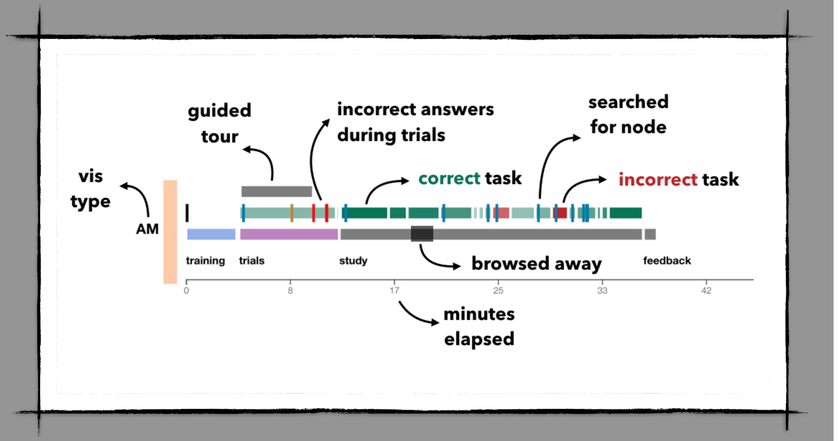


Visual Misinformation



EMPIRICAL & THEORETICAL WORK

Evaluating Interactive Visualizations



Zach Cutler, Jack Wilburn, Carolina Nobre, Lane Harrison, Hilson Shresta, Yiren Ding et al.

reVISit: Empirically Evaluating Complex Interactive Visualization Techniques

[CHI 2020, CHI 2021 & VIS 2023]

PERILS OF ONLINE VISUALIZATION EXPERIMENTS

There is no **end-to-end platform** for building **interactive experiments**

The support for **types of stimuli** in survey tools are limited

Lack of experiment **debugging** features

No built-in **data collection** and **provenance tracking** functions

A TOOLKIT FOR DATA VISUALIZATION EXPERIMENTS

visualization design lab

view

HIVE
Human-Interactive Visualization Lab

THE UNIVERSITY OF UTAH

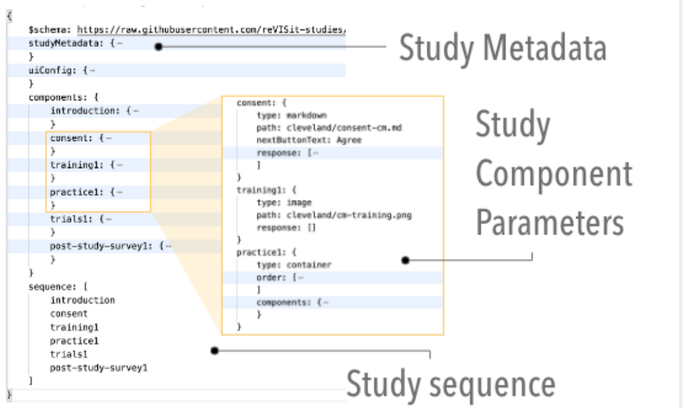
WPI

UNIVERSITY OF TORONTO

The Scope of reVISit

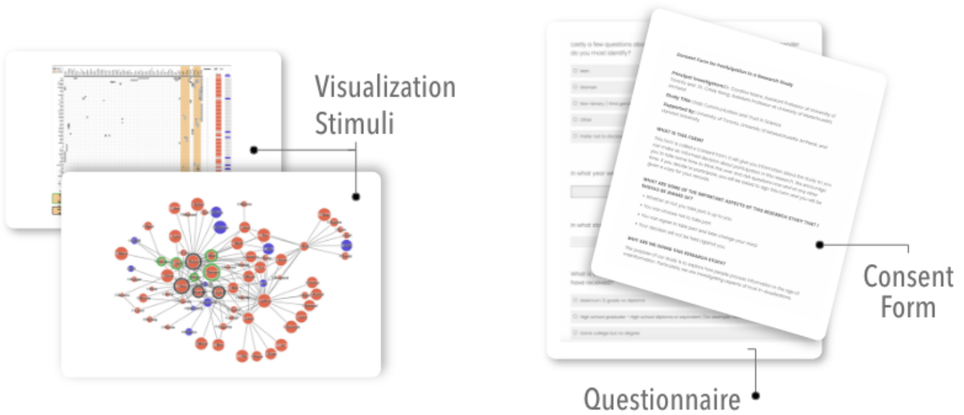
reVISit DSL - study specification

1



Study Stimuli & Components

2



Compile & Deploy

Interactive Study Web App

3

Task Instructions | Study Stimulus | Admin Panel

Data Management

4

Data Storage (Firebase)

Data Download (tidy text/csv, JSON)

Data Analysis

5

Study Analytics Interface

[reVISit CHI'21] & Future Work

External Data Analysis

R SPSS

WHAT IS REVISIT: INFRASTRUCTURE FOR VIS STUDIES

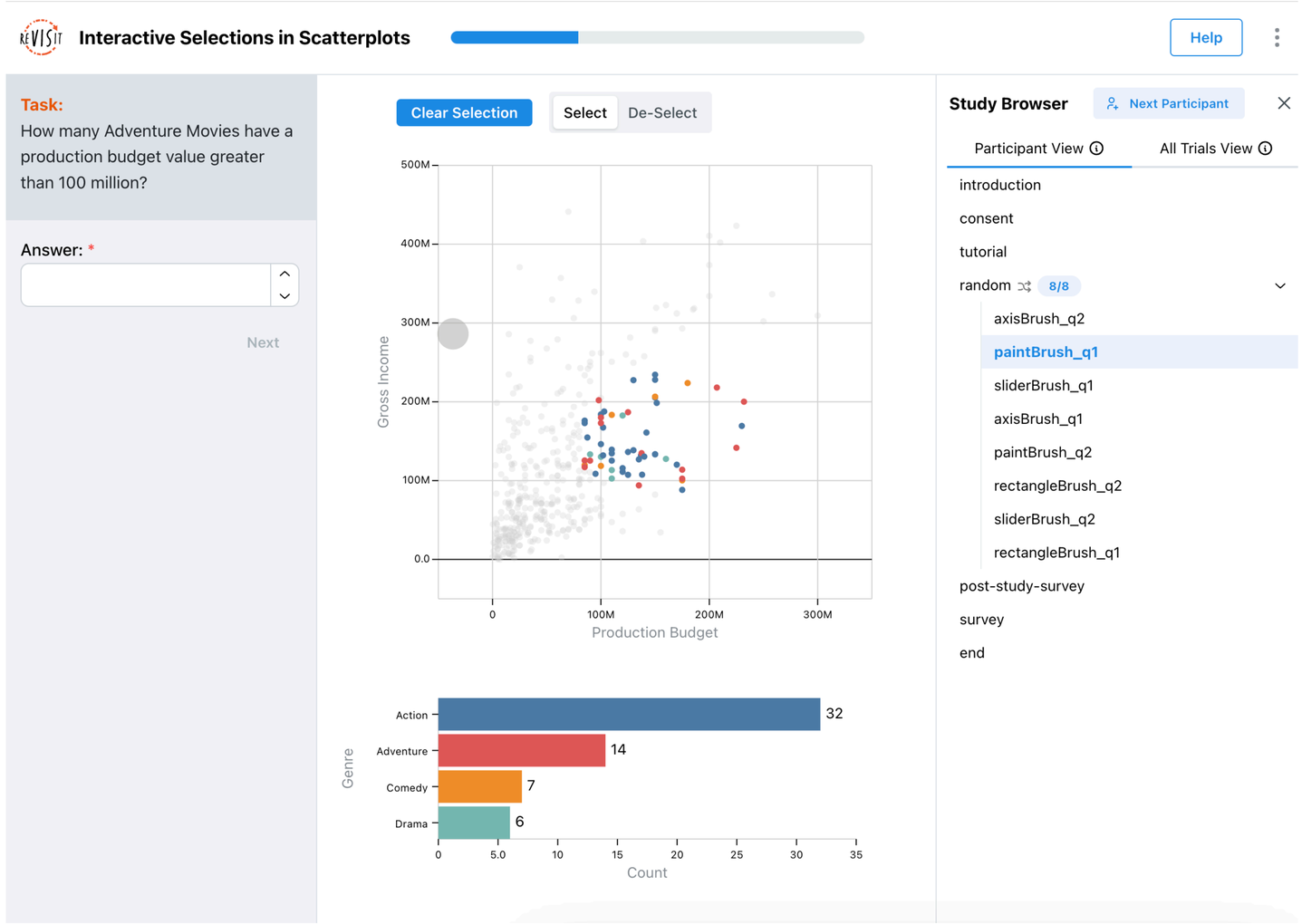
Set up studies with all **components** (consent, training, trials, tasks, surveys, etc.) including sophisticated **study designs** (randomization)

Simple **data tracking and data export**

Simple to **deploy and run** studies

Fully **reproducible and open**. You can share your whole study setup for anyone to re-run, without having to have a license for software (qualtrics).

DEMO



https://revisit.dev/study/demo-brush-interactions/

WHY REVISIT?

Designed for **sophisticated studies**

randomization, complex stimuli, complex data tracking

Reproducible and Open Studies!

Non-commercial license lets you **share your study!**

Extensible – add on features you need

Code not GUI! (though there might be a GUI option in the future)

Support from us!

THE REVISIT SPEC

Define the details of your experiment as a JSON file.

Study Metadata — name of the study, authors, contact e-mails

UI Config — parameterizing the appearance of reVISit

Components and **BaseComponents** — setting up the content of the study

Sequence — choosing the order and the selection of tasks participants see.

REVISIT SPEC

Outline

```
{
    "$schema":  "https://raw.githubusercontent.com/reVISit-studies/study/main/src/parser/
StudyConfigSchema.json",
    "studyMetadata": {
        "title": "Basic Questionnaire Study",
…
    },
    "uiConfig": {
        "contactEmail": "contact@revisit.dev",
…
    },
    "components": {
        "introduction": {
            "type": "markdown",
            "path": "basic-questionnaire-study/assets/introduction.md",
            "response": []
        },
        "first-question-set": {
            "type": "questionnaire",
            "response": [
                {
                    "id": "q1-name",
                    "prompt": "What is your first name?",
                    "type": "longText",
                    "placeholder": "Please enter your first name"
                },
    "sequence": {
        "order": "fixed",
        "components": [
            "introduction",
            "first-question-set"]
    }
```

REVISIT COMPONENTS – WHERE YOUR "STUDY CONTENT" GOES

Markdown Files – introductions, consent forms, help pages, etc.

Images – static (vis) stimuli

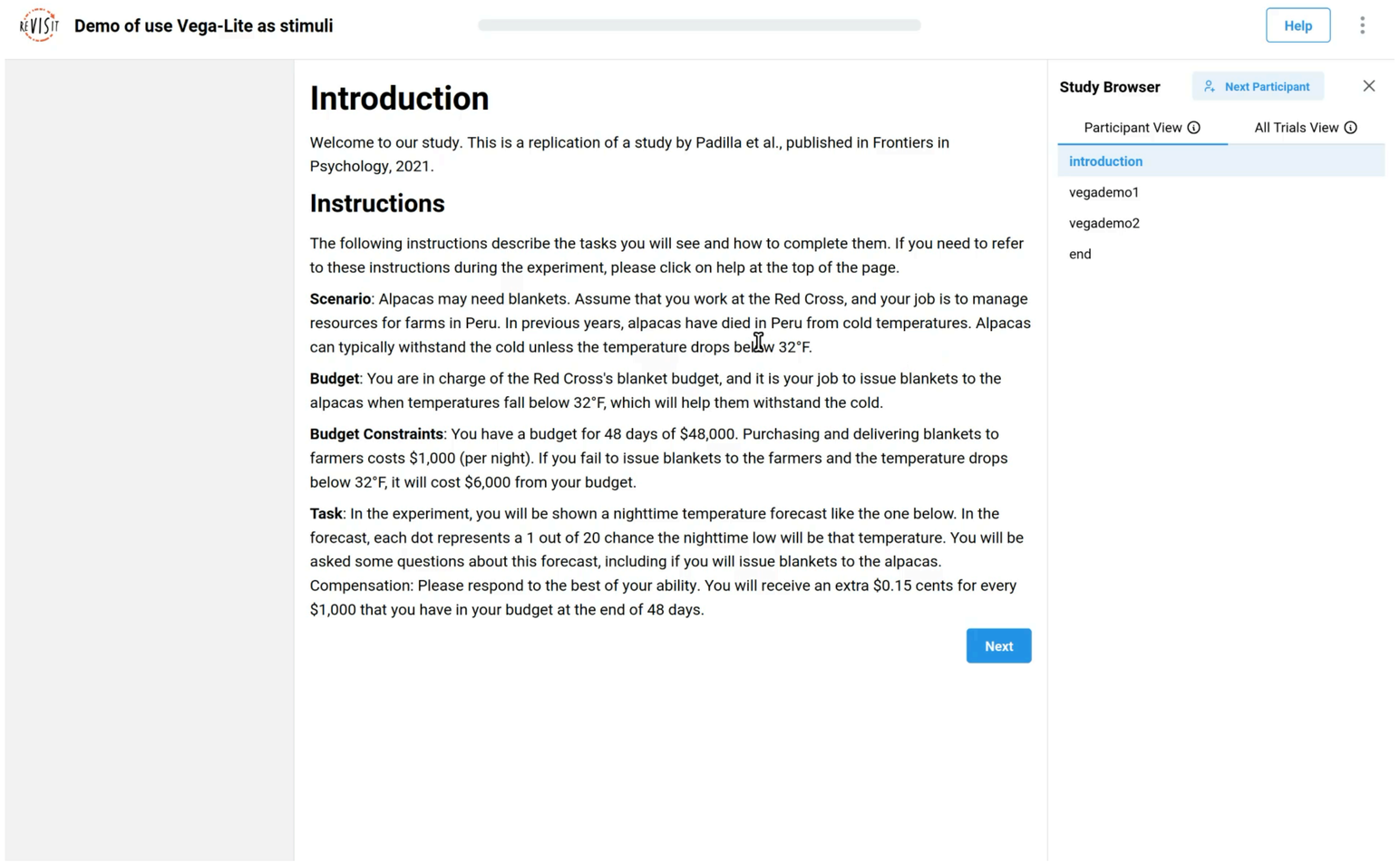
Web Pages – custom stimuli, made interactive with JavaScript

React Components – sophisticated interactive stimuli. Simplify the communication between reVISit and the stimulus

Survey Questions – structured responses

Vega Visualizations – directly put in interactive visualizations as stimulus (beta)

VEGA DEMO



Also auto-captures interaction!

REVISIT SPEC

Components & Inheritance

```
    "baseComponents": {
        "bar-chart": {
            "type": "website",
            "response": [
                {
                    "id": "barChart",
                    "prompt": "Your selected answer:",
                    "required": true,
                    "location": "belowStimulus",
                    "type": "iframe"
                }
            ],
            "path": "basic-questionnaire-study/assets/bar-chart.html",
            "instructionLocation": "aboveStimulus"

        }
    }

"bar-chart-1":{
    "baseComponent": "bar-chart",
    "description": "A trial for the user to click the largest bar",
    "instruction": "Click on the largest bar",
    "parameters": {
        "barData": [0.32, 0.01, 1.2, 1.3, 0.82, 0.4, 0.3]
    }
},
"bar-chart-2":{
    "baseComponent": "bar-chart",
    "description": "A trial for the user to click the smallest bar",
    "instruction": "Click on the smallest bar",
    "parameters": {
        "barData": [1.2, 1.2, 1.2, 1.3, 0.82, 0.4, 0.3]
    }
}
```

COMPONENTS AND RESPONSES

Responses are a **primary data type – each component can have a response**

Form elements are “responses only”

Component

Response

```
"first-question-set": {
                "type": "questionnaire",
                "response": [
                    {
                        "id": "q1-name",
                        "prompt": "What is your first name?",
                        "required": true,
                        "location": "aboveStimulus",
                        "type": "longText",
                        "placeholder": "Please enter your first name"
                    },
```

SEQUENCE

In what order do the components appear for which participant

- Can be made up of **nested blocks**
- **Fixed**
- **Random**
- **Latin Square**
- **Subsets** (show 2 out of 5)
- **Skips** (if wrong, go to next task)
- **Interrupts** (breaks, attention checks)

REVISIT SPEC

Sequence

```
"sequence": {
        "order": "fixed",
        "components": [
            "introduction",
            "consent",
            "tutorial",
            {
                "order": "random",
                "components": [
                    "paintBrush_q1",
                    "rectangleBrush_q1",
                    "axisBrush_q1",
                    "sliderBrush_q1"
                ]
            },
            "post-study-survey",
            "survey"
        ]
    }
```

DATA COLLECTION

Responses – Participant responses to specified questions

Logs – a variety of things about participants and sessions

- **time** information about responses
- **“Clean” time** - minus “browsed away” time
- **Participant Metadata**: Browser information, etc

Provenance Data – data you track in your stimulus via our provenance tracking library

Audio Data – you can run think aloud studies (more later)

PROVENANCE DATA

Instrument your interactive stimulus with provenance tracking!

E.g., using our **trrack library**

Works out **of the box for Vega!**

https://apps.vdl.sci.utah.edu/trrack

Provenance **data is stored as JSON** with the rest of your data

PROVENANCE DATA

New analysis possibilities!

- Analyze **logs**
- Analyze **interaction strategies**

Review individual user actions

- **Debug** your study setup
- See where things go **wrong**

Review Demo

DATA STORAGE: LOCAL OR FIREBASE

Local

- Works out of the box
- Download from local storage of your browser
- Doesn’t work for online studies

Firebase

- A bit of a pain to set up (but better than running your own server)
- Reliable cloud storage option
- Available with servers in many countries

STUDY MODES — EXPERIMENT LIFECYCLE

Study Design & Debugging

Jump around a study, see all data, see run-throughs of pilots

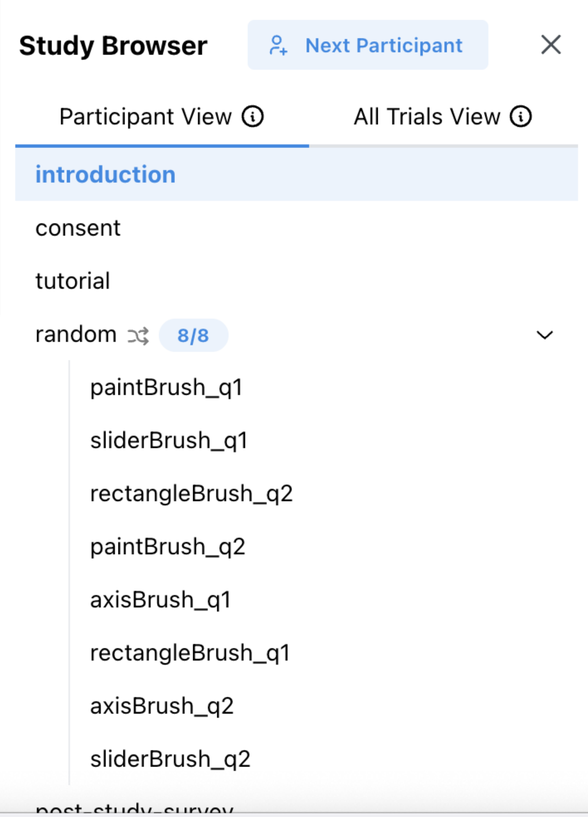
Collecting Participant Data

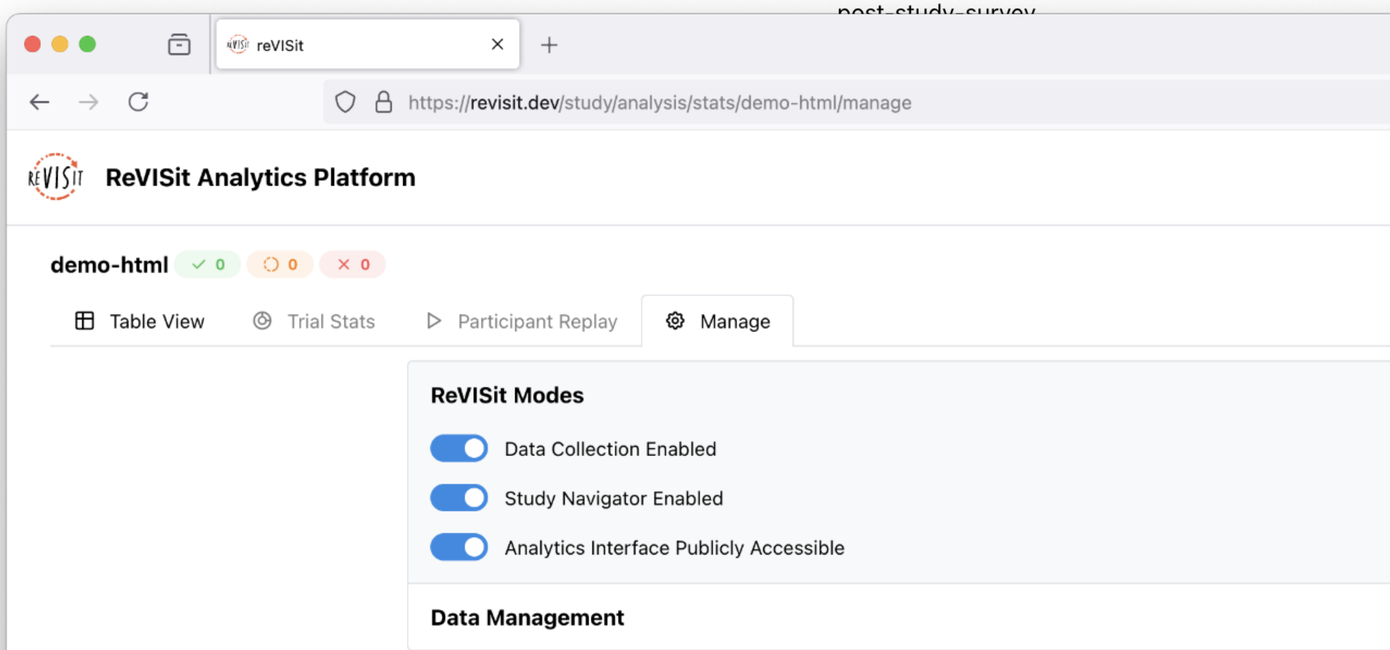
Hide all debug features

Study & Data Dissemination

Share study w. navigation

Share data

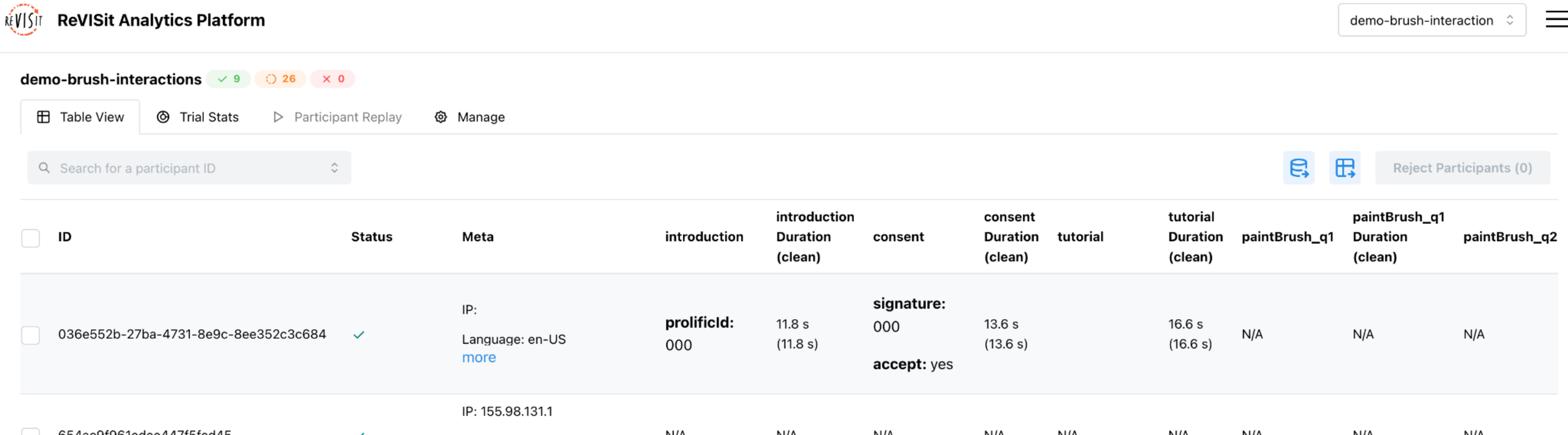




ANALYSIS INTERFACE

Download the data as tidy csv or JSON

Review and reject trials

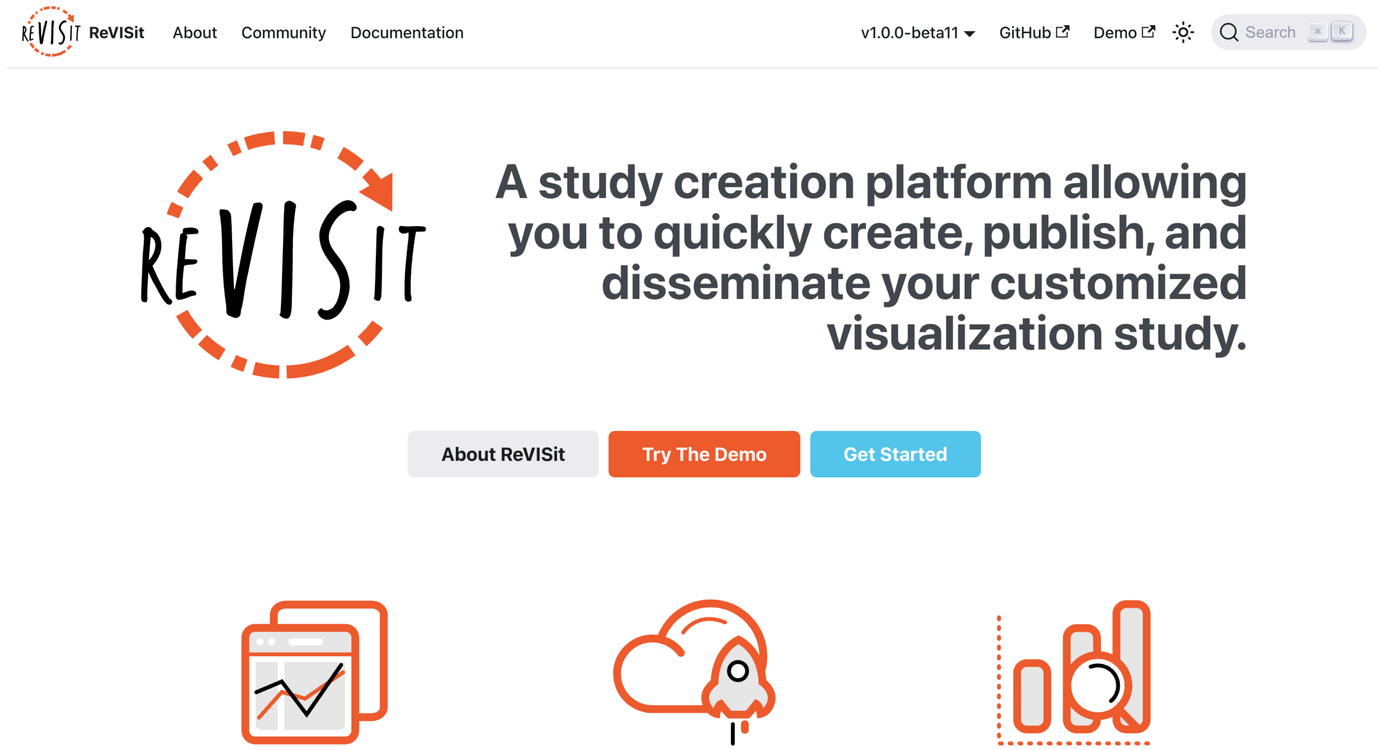


DOCUMENTATION & COMMUNITY

Documentation & Tutorials on Website

Community participation via slack etc.

https://revisit.dev



FUTURE DIRECTIONS

Looking for community input!

Improve Analysis Interface

Sequencing: step-functions

Integration with prolific: balancing latin squares

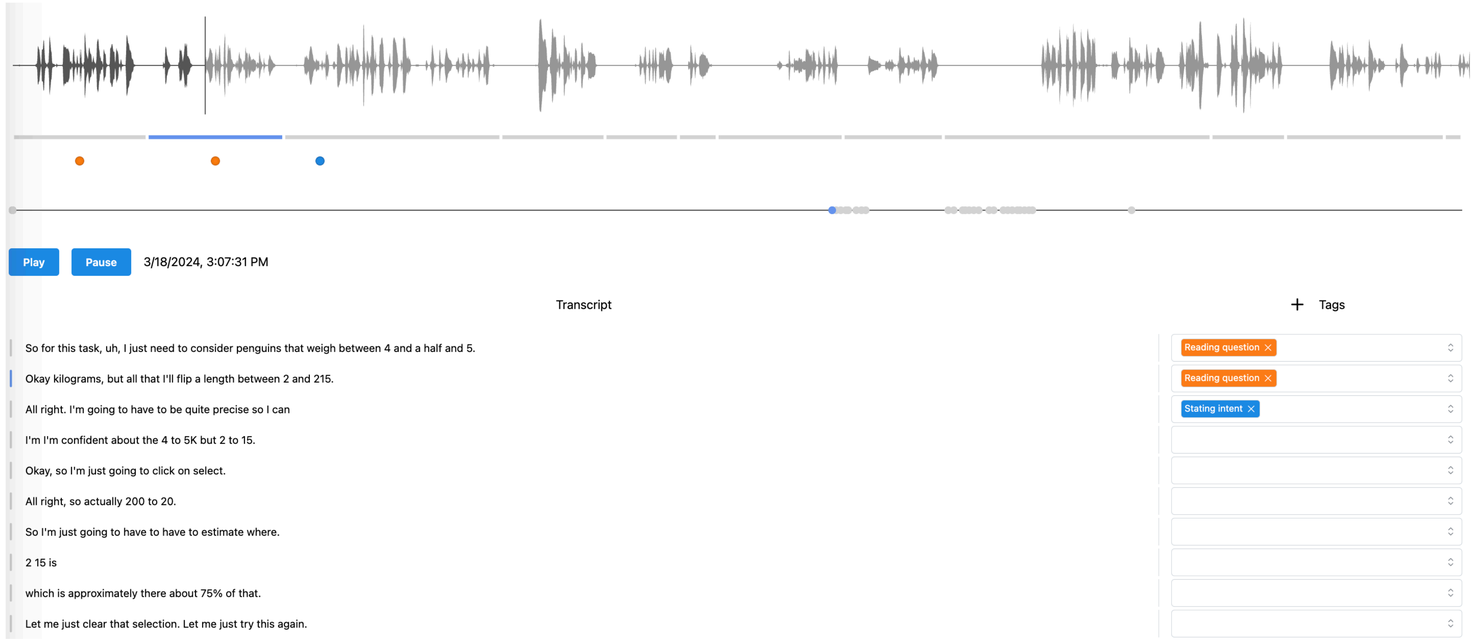
Libraries of common surveys / tasks

Python Bindings

Higher-level language features

CAN WE DO QUALITATIVE EVALUATION ONLINE?

Yes! Think Aloud & Provenance



Pushing the boundary of what can be evaluated using crowdsourcing

THINK ALOUD PROTOCOL

A UI/UX Method for finding usability issues in software

Participants sit in a **lab**

Instructed to "**speak their thoughts**" as they use an interface

Audio is recorded

Screen is recorded

Logs may be recorded

Experimenter is **present**

PURPOSE

Traditional focus on “**Defects**” and UI Issues

Vis community also uses it to analyze **Insight Formation** with visualizations

REVIEWED 67 TA STUDIES IN VIS

66 were synchronous (lab or zoom)

27 evaluated insights

Participants:

- **experts** – evaluating a domain-specific system (small n)
- **novices & skilled participants** (large n, 20-50)

RESEARCH QUESTIONS

Can we build a **system** that makes **recording and analyzing TA studies** online easy?

Do **crowdsourced TA studies** actually **work**?

WHY?

Logistics are easier

- Crowdsourcing
- Expert-studies at their schedule

Potentially more **diverse participant pools**

Time-saving for experimenter

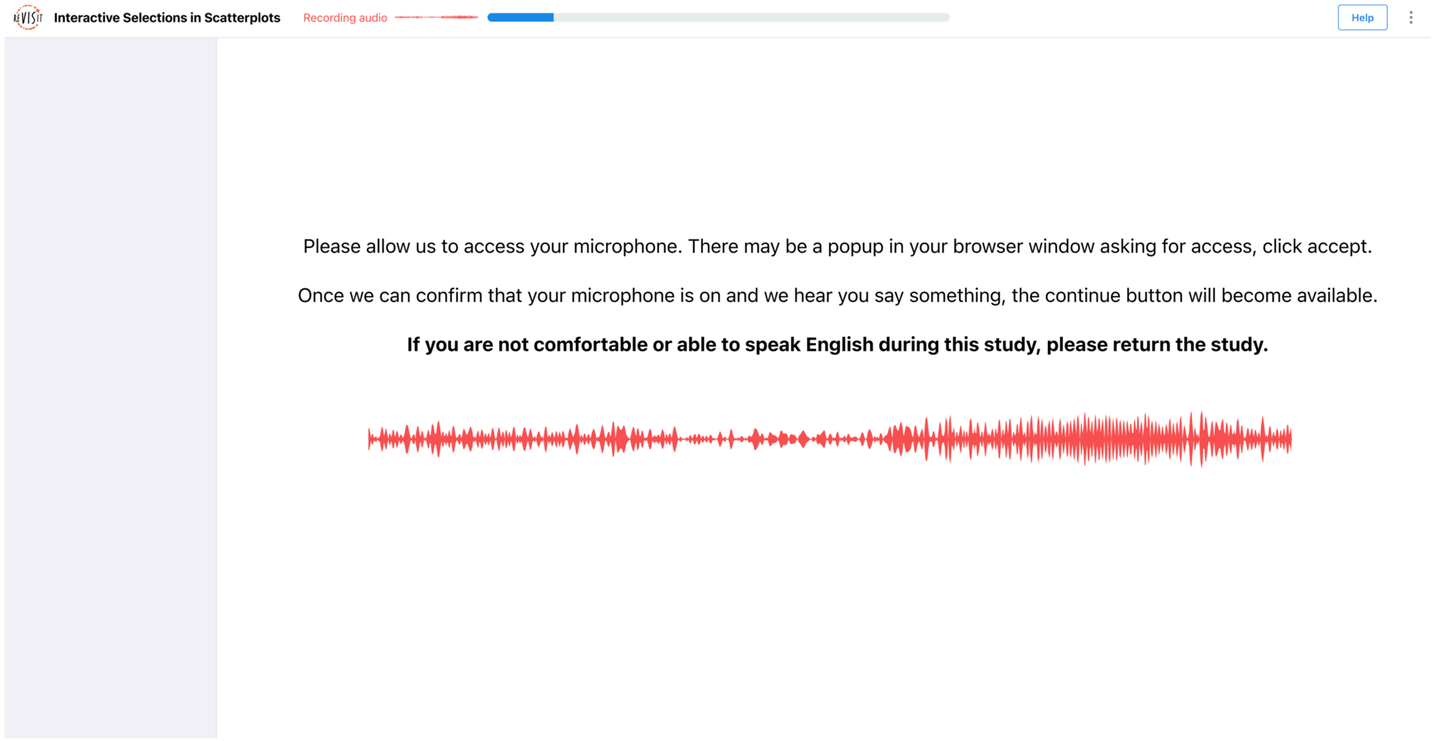
Usual trade-offs of lab vs crowdsourced

RECORDING

Single flag

Waveform vis to show recording is on

Automatic **transcription**



ANALYSIS

Synched with Application State / Events

Easy **editing** of transcripts

Ability to **code** per participant / task / trial

Interactive Selections in Scatterplots

Help

5e5521580ee1b951df544c3c

paintBrush_q4

Play Pause 3/18/2024, 3:08:05 PM

Transcript	Tags
So for this task, uh, I just need to consider penguins that weigh between 4 and a half and 5.	Reading question
Okay kilograms, but all that I'll flip a length between 2 and 215.	Reading question
All right. I'm going to have to be quite precise so I can	
I'm I'm confident about the 4 to 5K but 2 to 15.	
Okay, so I'm just going to click on select.	Stating Intent
All right, so actually 200 to 20.	
So I'm just going to have to have to estimate where.	
2 15 is	
which is approximately there about 75% of that.	
Let me just clear that selection. Let me just try this again.	
Okay. Alright, so that's the 4 to 5K body mass. And then the 2002215.	
So that's giving me those results.	
so the most common penguin There Is the Gen 2	Answering question
and the least common penguin with 5 is the a deli if I've pronounced that correctly.	
Okay clicking.	

Interactive Selections in Scatterplots

Help

Task:

Consider only penguins that weigh between 4k and 5k grams (g) and have flipper lengths between 200 and 215mm. What is the most and least common type of penguin in this subset?

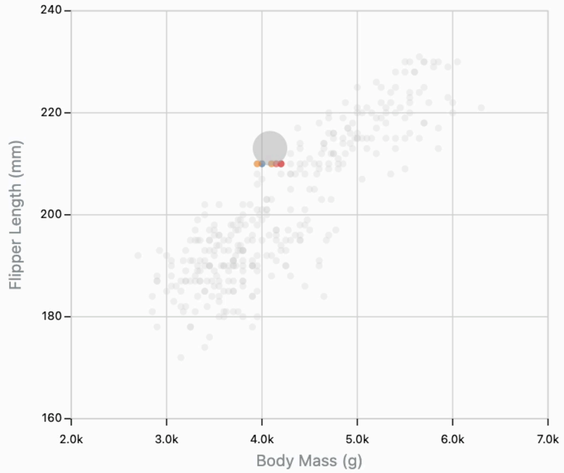
While answering this question, please verbalize your thoughts, especially any insights you have or problems you run into.

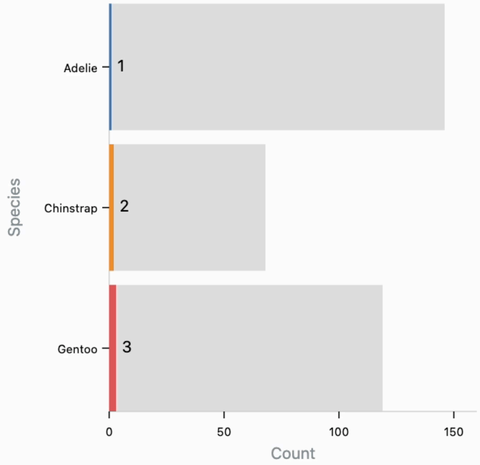
Most common Penguin: *

Least common Penguin: *

Next

Clear Selection Select De-Select





VALIDATION

Do crowdsourced think-aloud studies work?

Study 1: Crowdsourced vs Lab Think-Aloud

Crowdsourced

Lab

Study 2: Online Think-Aloud vs Text Response

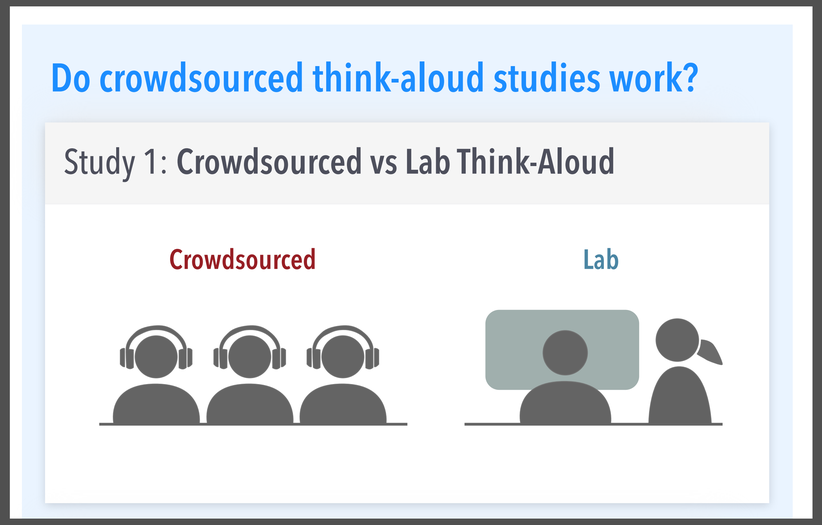
Think-Aloud

RECORD RESPONSE

Text Response

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut ero labore et dolore

SUBMIT RESPONSE



40 online participants

10 lab participants

Stimulus:

interactive vis tool

two views

required interactions to solve tasks

Coding:

Usability

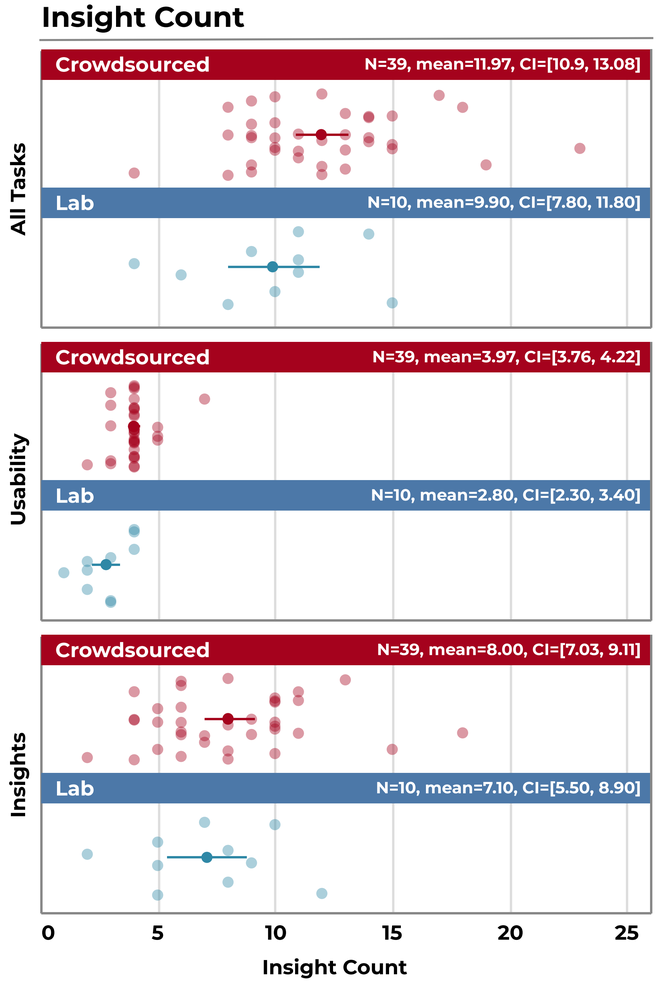
Insights (correct, directed / open, hypothesis, elaborated)…

Do crowdsourced think-aloud studies work?

Study 1: **Crowdsourced vs Lab Think-Aloud**

Crowdsourced

Lab

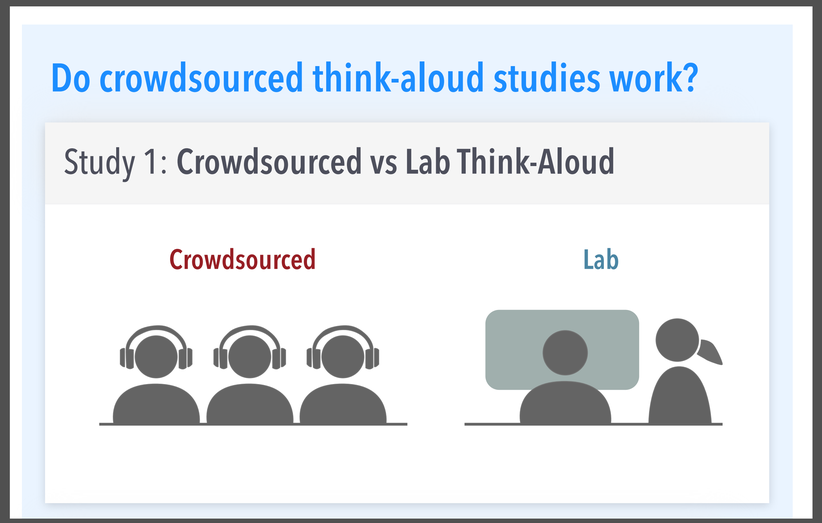


Yes, ppl speak!

1 audio recording unusable

2 required heavy editing

recruiting slightly slower



Crowdsourced comments were more negative

- Bias because of experimenter present in lab setting?

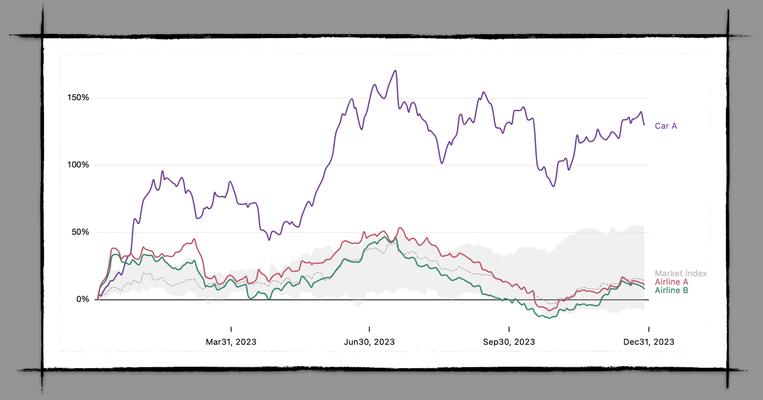
Role of experimenter

- Better introduction
- Clarifying questions

Need to carefully prepare crowdsourced study!

EMPIRICAL & THEORETICAL WORK

Max Lisnic, Marina Kogan

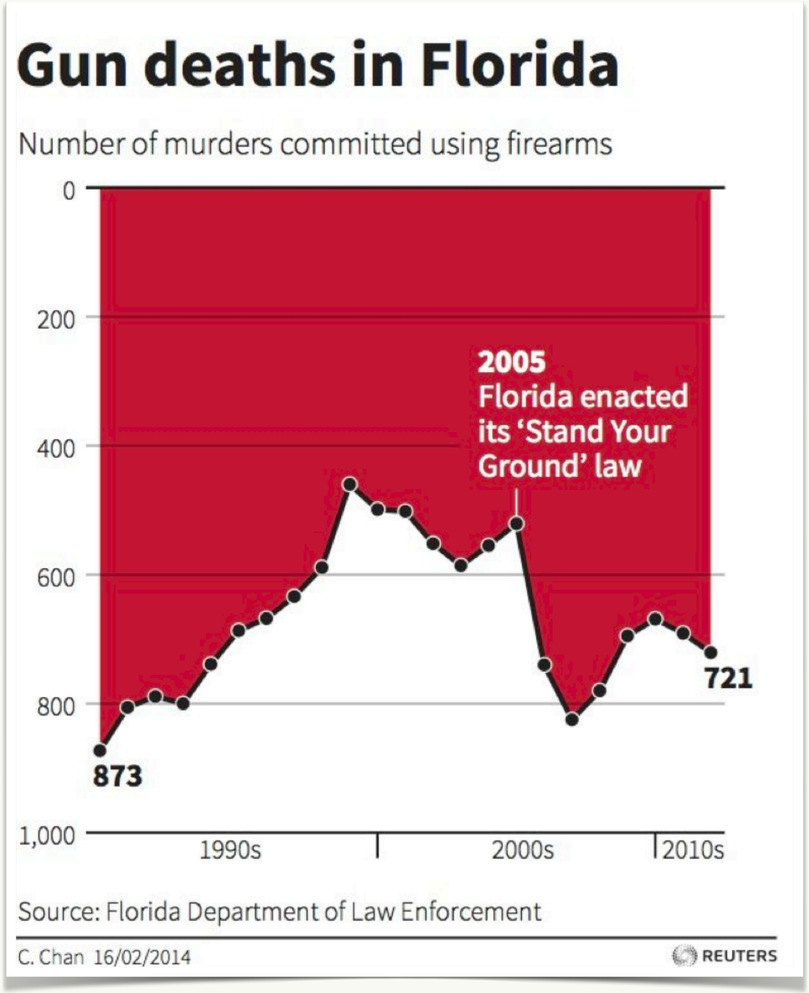


HOW PEOPLE ACTUALLY LIE WITH CHARTS. . .

. . . AND WHAT WE CAN DO ABOUT IT

[CHI 2023 & CHI 2024]

HOW WE THINK PPL LIE WITH CHARTS: VIOLATIONS OF VISUALIZATION DESIGN GUIDELINES



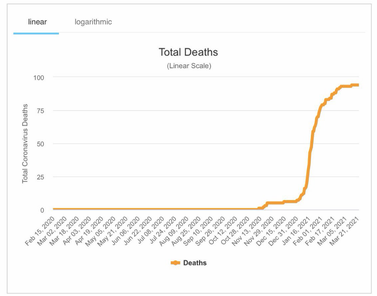
Inverted y axis

Truncated y axis

BUT IS THAT REALLY HOW PEOPLE LIE WITH CHARS?

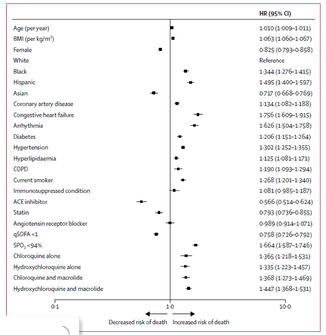
As of 10th January when they started vaccinations there had only been 16 cases.

Total Coronavirus Deaths in Gibraltar

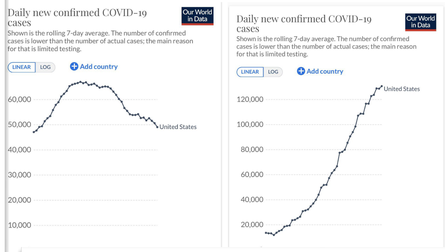


Taking HCQ is as strongly associated with increased coronavirus death risk as DIABETES.

thelancet.com/journals/lance...

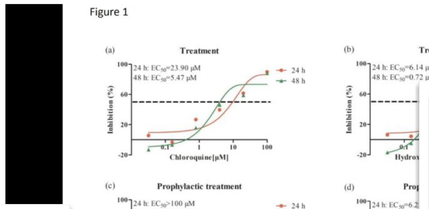


New daily covid cases this time last year (pre-vaccination) vs this year (post-vaccination)



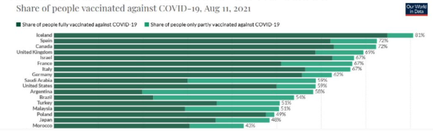
Hydroxychloroquine probably better

m.youtube.com/watch?feature=...

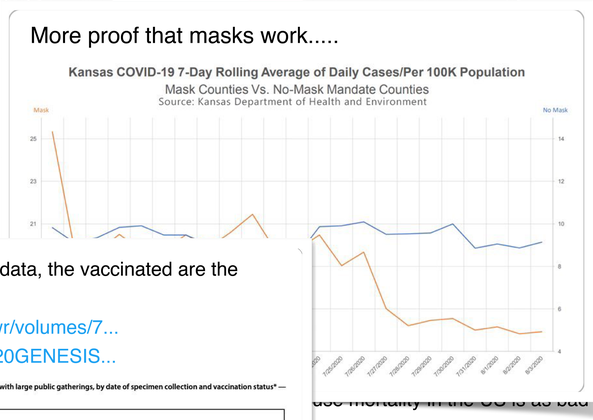


Most vaccinated: Iceland 81%
Least vaccinated: Nigeria 1.2%

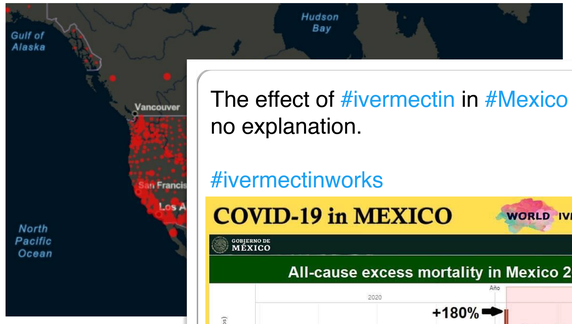
Iceland has 119 times more Covid cases



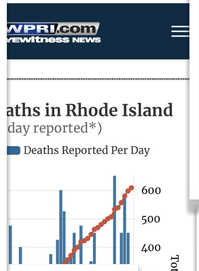
More proof that masks work.....



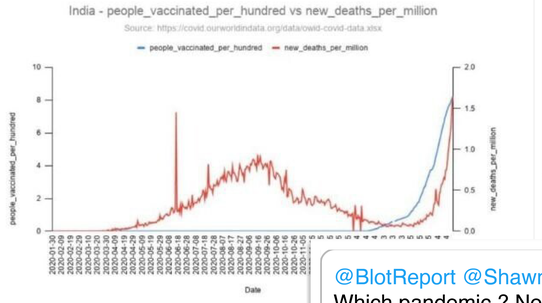
We need national leadership...like Canada.
Covid is coded in red.



milestone: today Rhode Island's
avirus death toll passed 600 peo

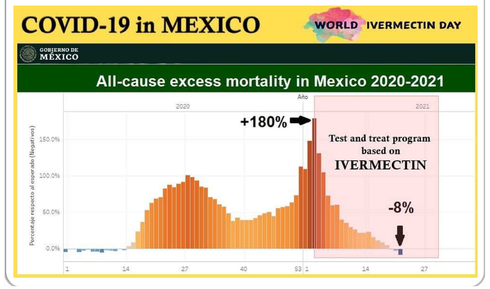


Whoa look what I found, death rates rising in line with vaccination numbers in India.



The effect of #ivermectin in #Mexico needs no explanation.

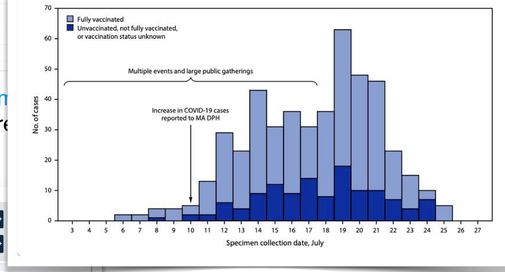
#ivermectinworks



According to official data, the vaccinated are the super-spreaders.

https://cdc.gov/mmwr/volumes/7...

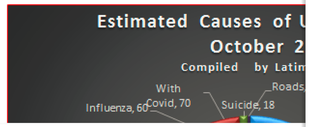
https://t.me/EARTH20GENESIS...



@BlotReport @ShawnG927 @ici_can

Which pandemic ? Nothing to see here

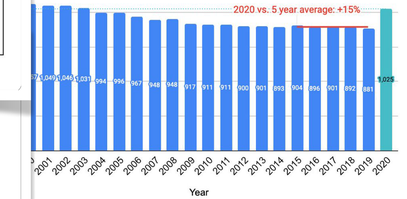
Hot off the press:
just 1 death in 23 today in UK wa



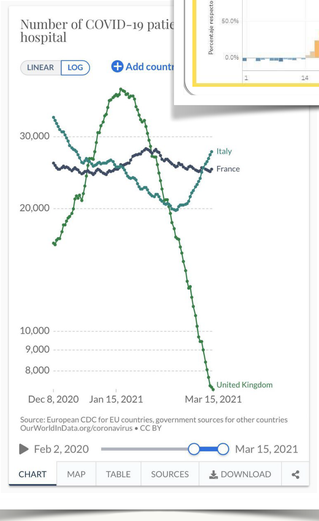
#Covid #Corona #Coronavirus

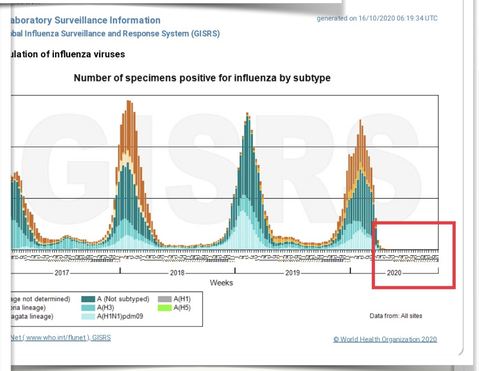
usted Mortality Rate [United States]

00; Adjusted for Age Distribution (United States 2020)

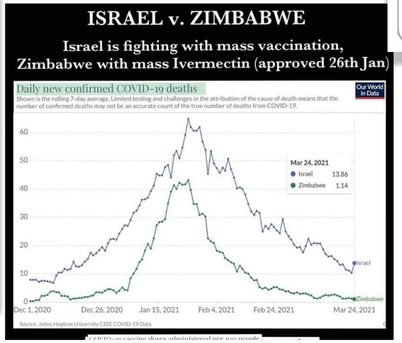


Vaccines work.

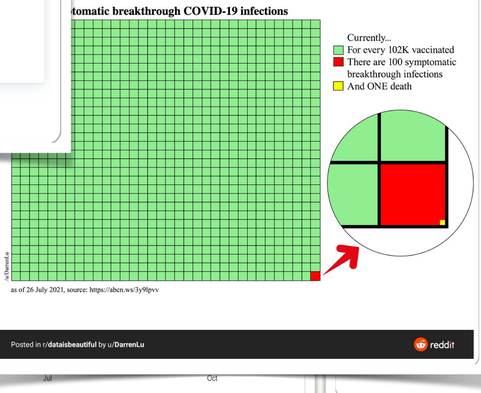


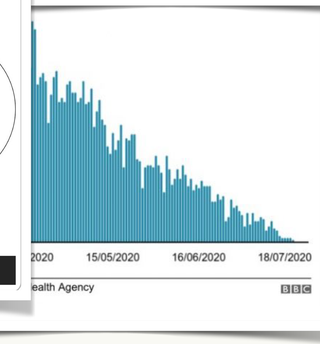


Two words
Ivermectin works
#ivermectinworks

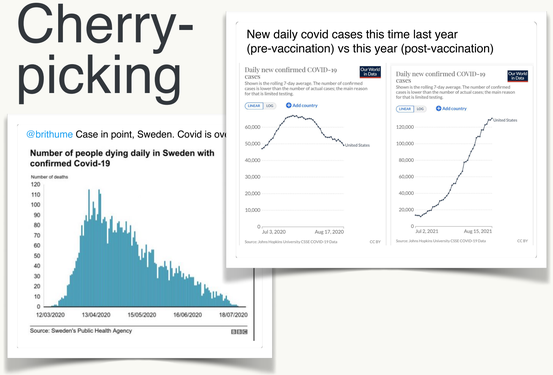


ccinated HT @FungalDoc & Reddit.

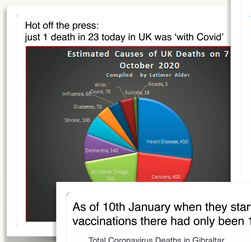


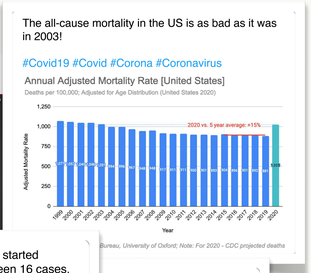


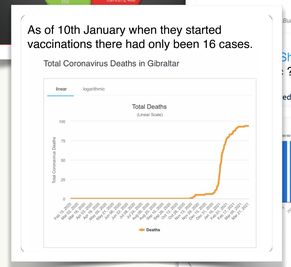
Cherry-picking



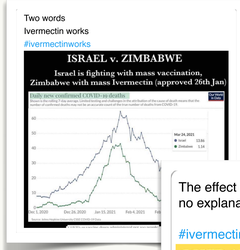
Arbitrary threshold

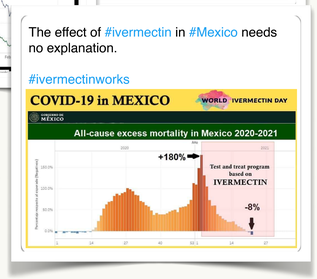




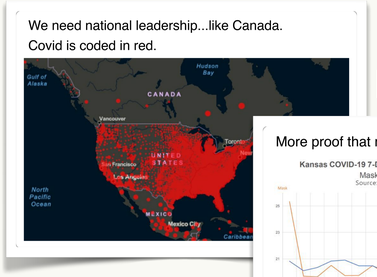


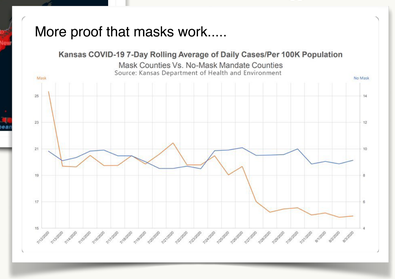
Causal inference



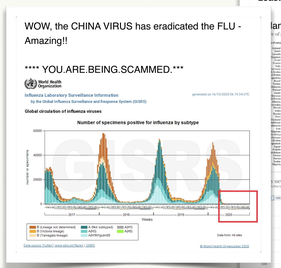


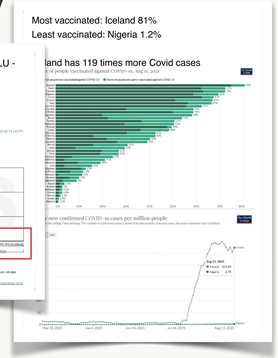
Incorrect reading



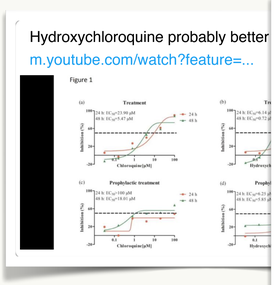


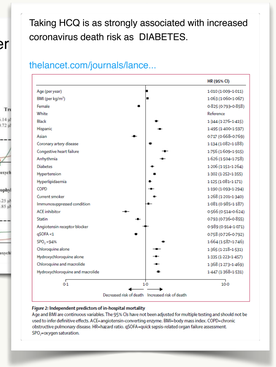
Issues with data validity



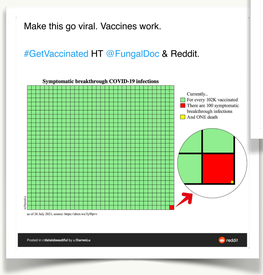


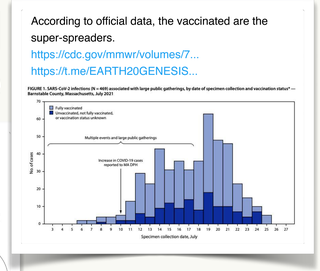
Misrepresentation of scientific results





Statistical nuance





The effect of #ivermectin in #Mexico needs no explanation.

#ivermectinworks

Causal inference

Screenshot of a government dashboard

Sharp mortality drop is a salient visual feature

Added annotation assigns a cause-and-effect

Tweet text further explains the argument

TAKEAWAYS

89% of charts

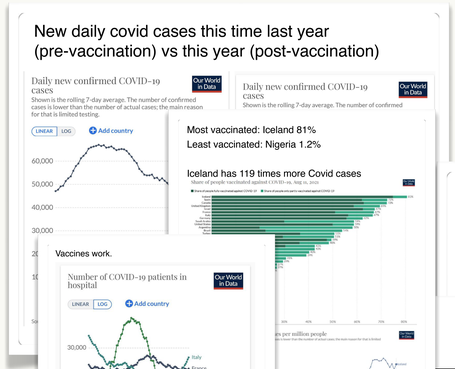
with reasoning errors

do not violate any visualization guidelines

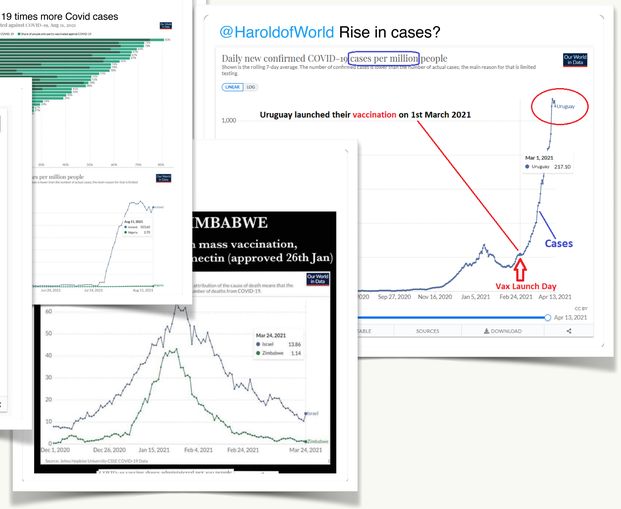
The majority of **misleading charts** are screenshots

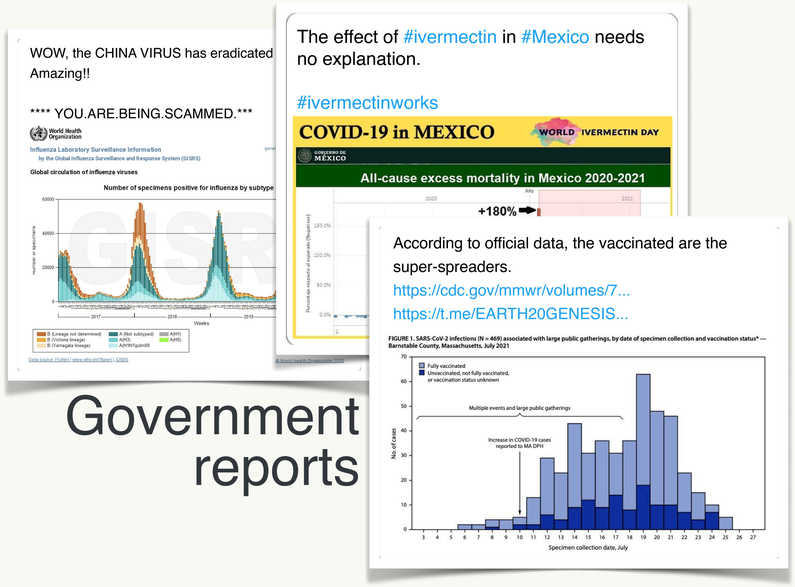
from **reputable sources**

Sources of Deceptive Visualizations

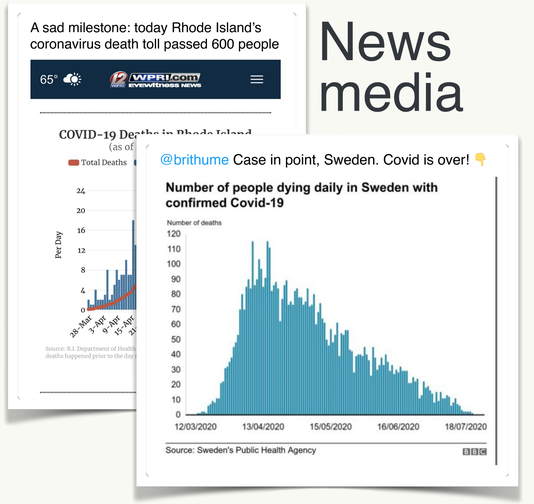


Data exploration websites





Government reports



News media

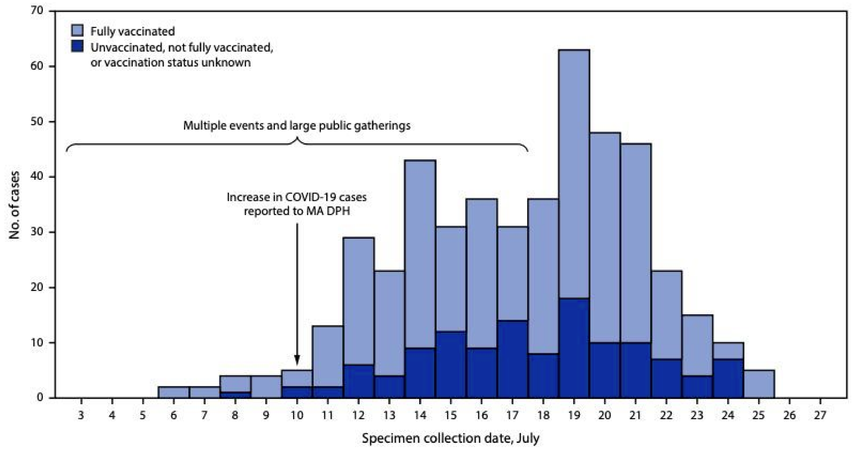
VULNERABLE VISUALIZATIONS

According to official data, the vaccinated are the super-spreaders.

https://cdc.gov/mmwr/volumes/7...

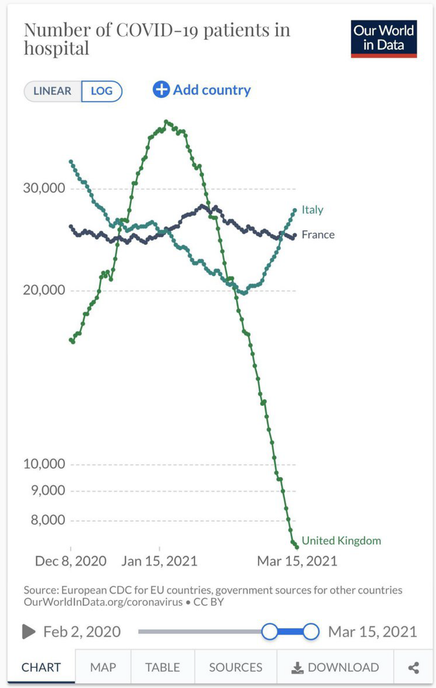
https://t.me/EARTH20GENESIS...

FIGURE 1. SARS-CoV-2 infections (N = 469) associated with large public gatherings, by date of specimen collection and vaccination status* — Barnstable County, Massachusetts, July 2021



Notes outside the chart

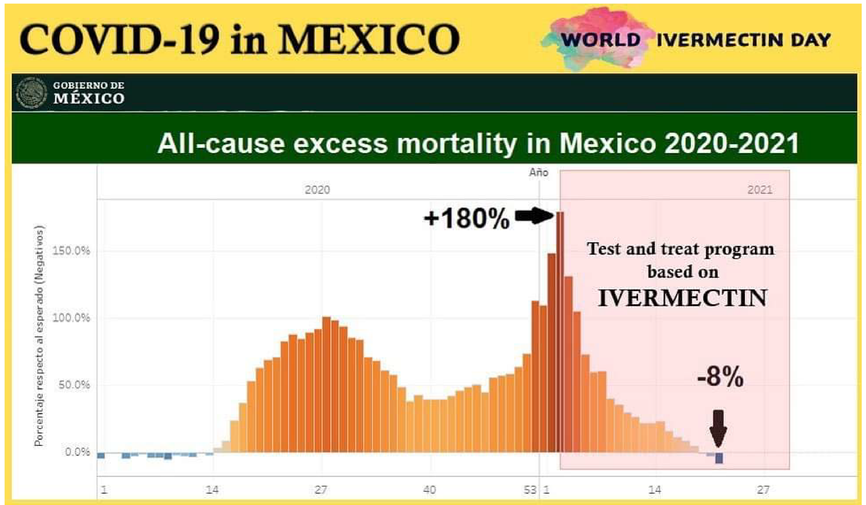
Vaccines work.



Unrestricted interaction

The effect of #ivermectin in #Mexico needs no explanation.

#ivermectinworks



Added annotations

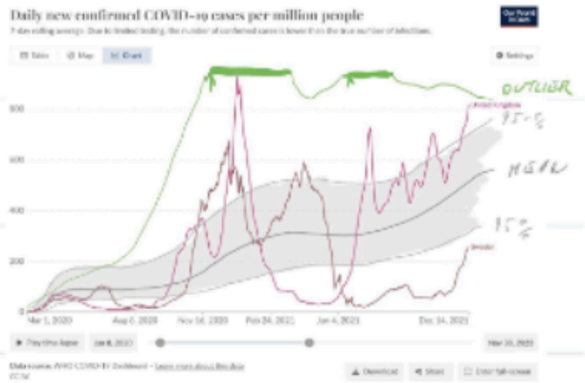
PROTECTING VISUALIZATIONS

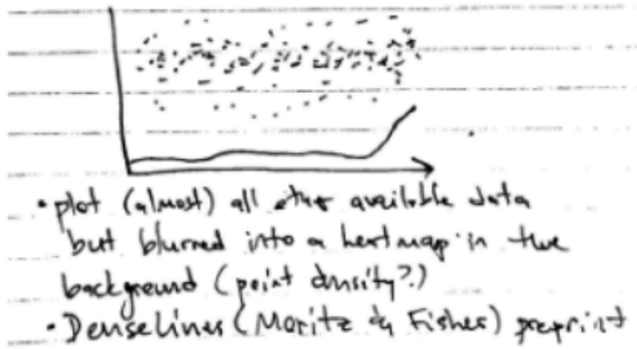
Can we design data exploration tools that prevent cherry-picking while maintaining freedom of exploration?

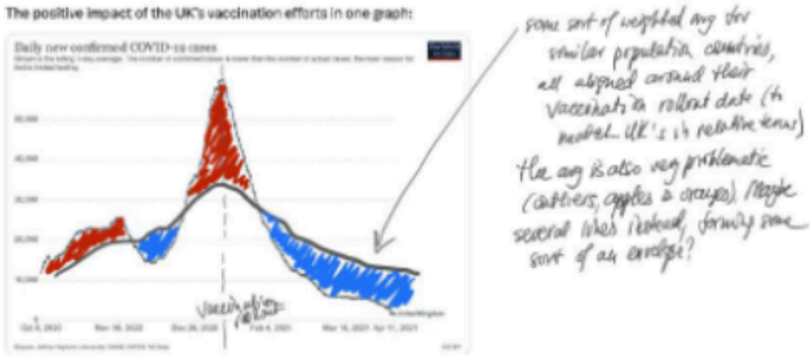
DESIGNING GUARDRAILS

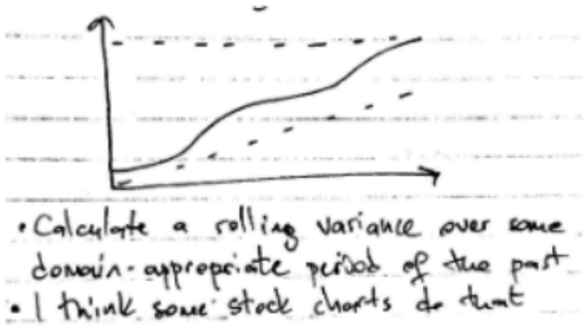
Parallel Design by 4 Authors

Cheery-Picking in Temporal Charts (Stocks, COVID)

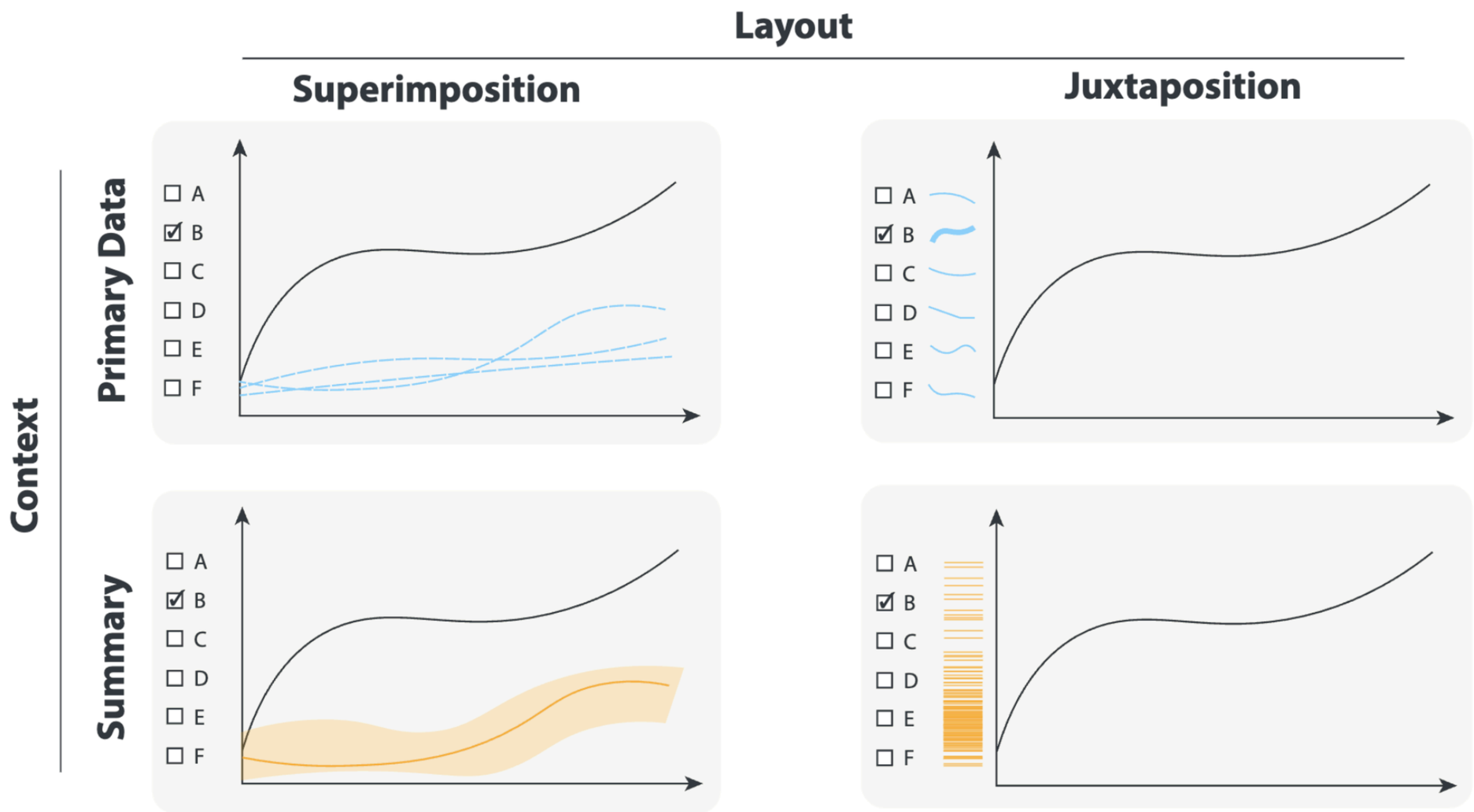




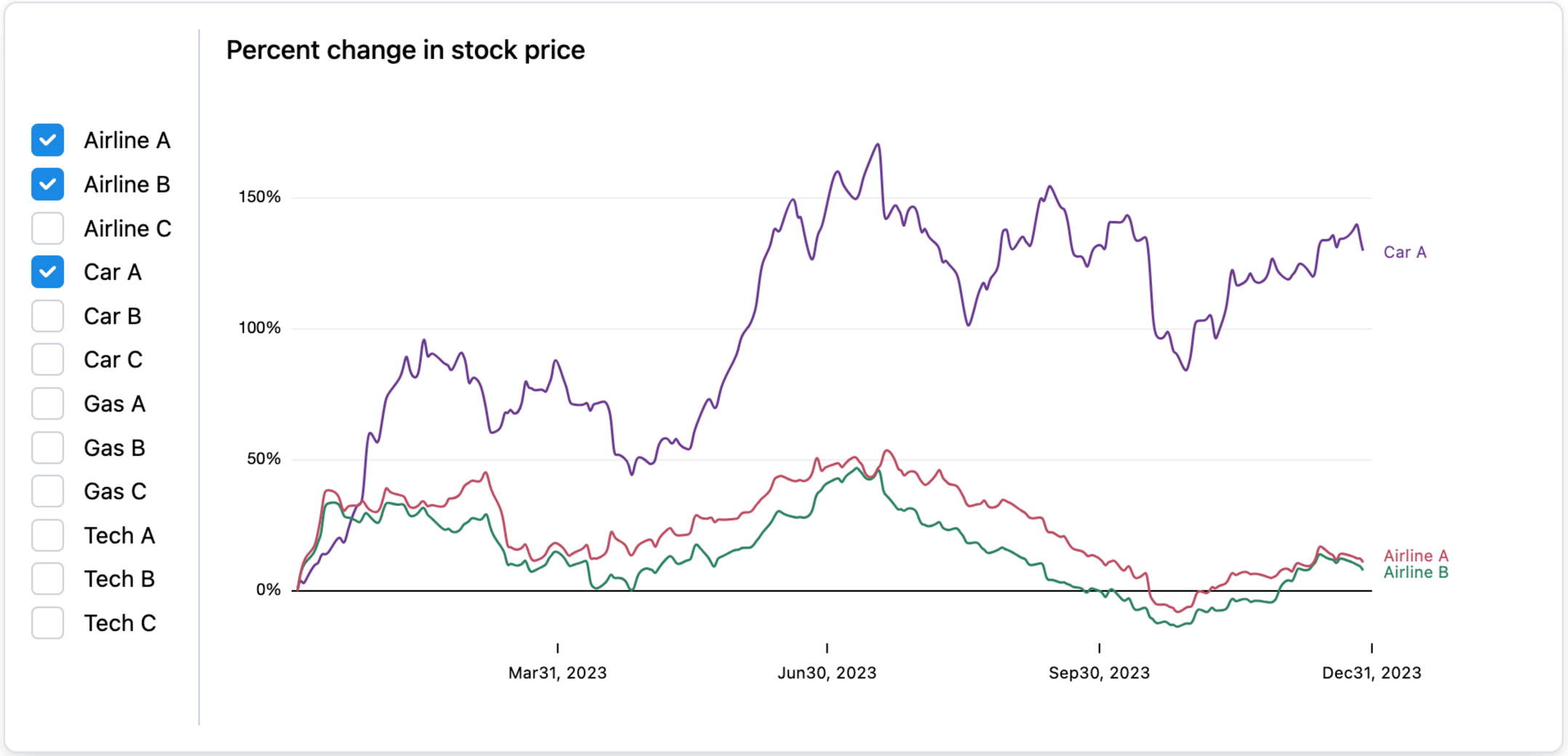




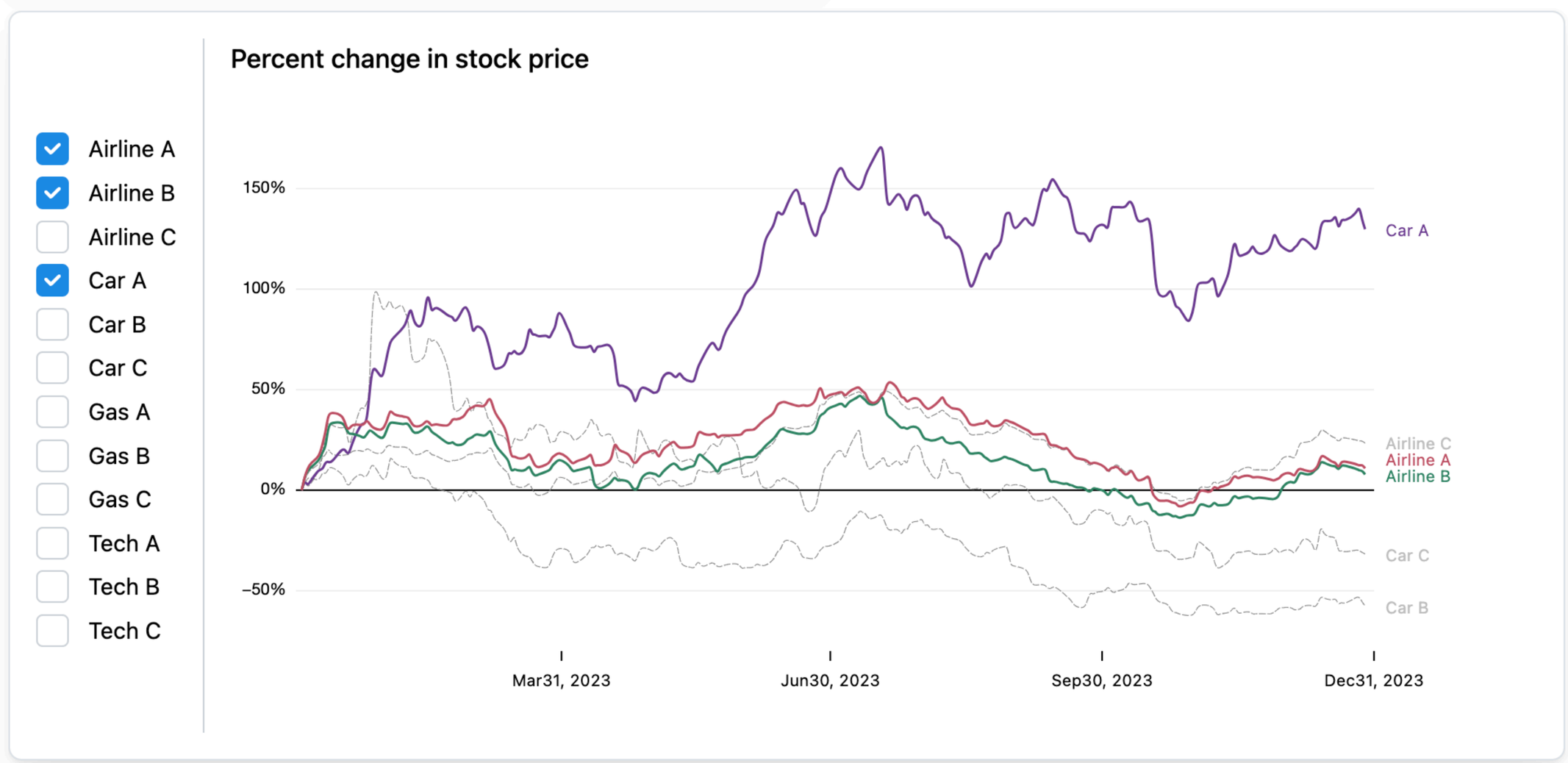
DESIGN SPACE



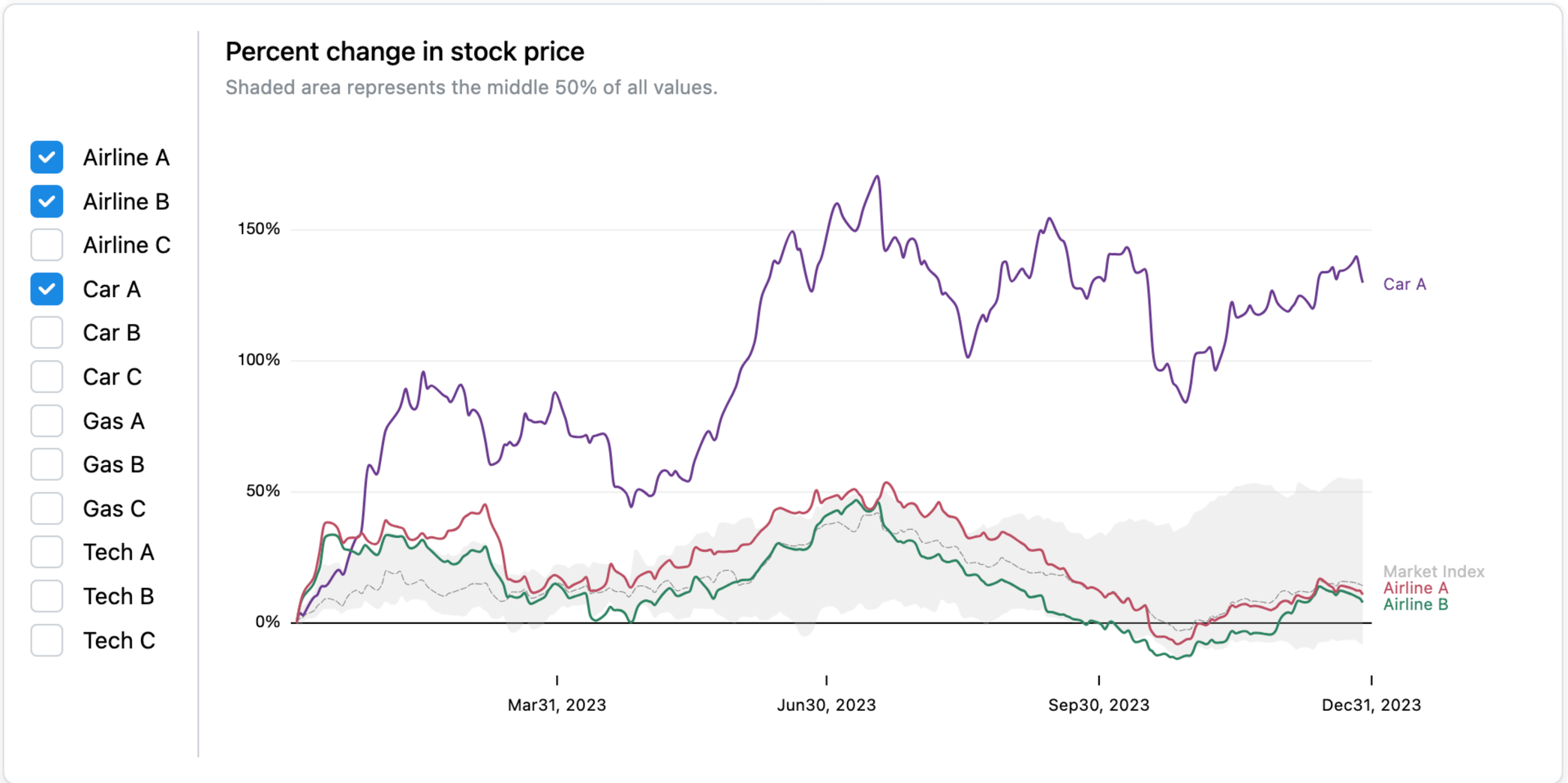
CHERRYPICKING: NO GUARDRAILS



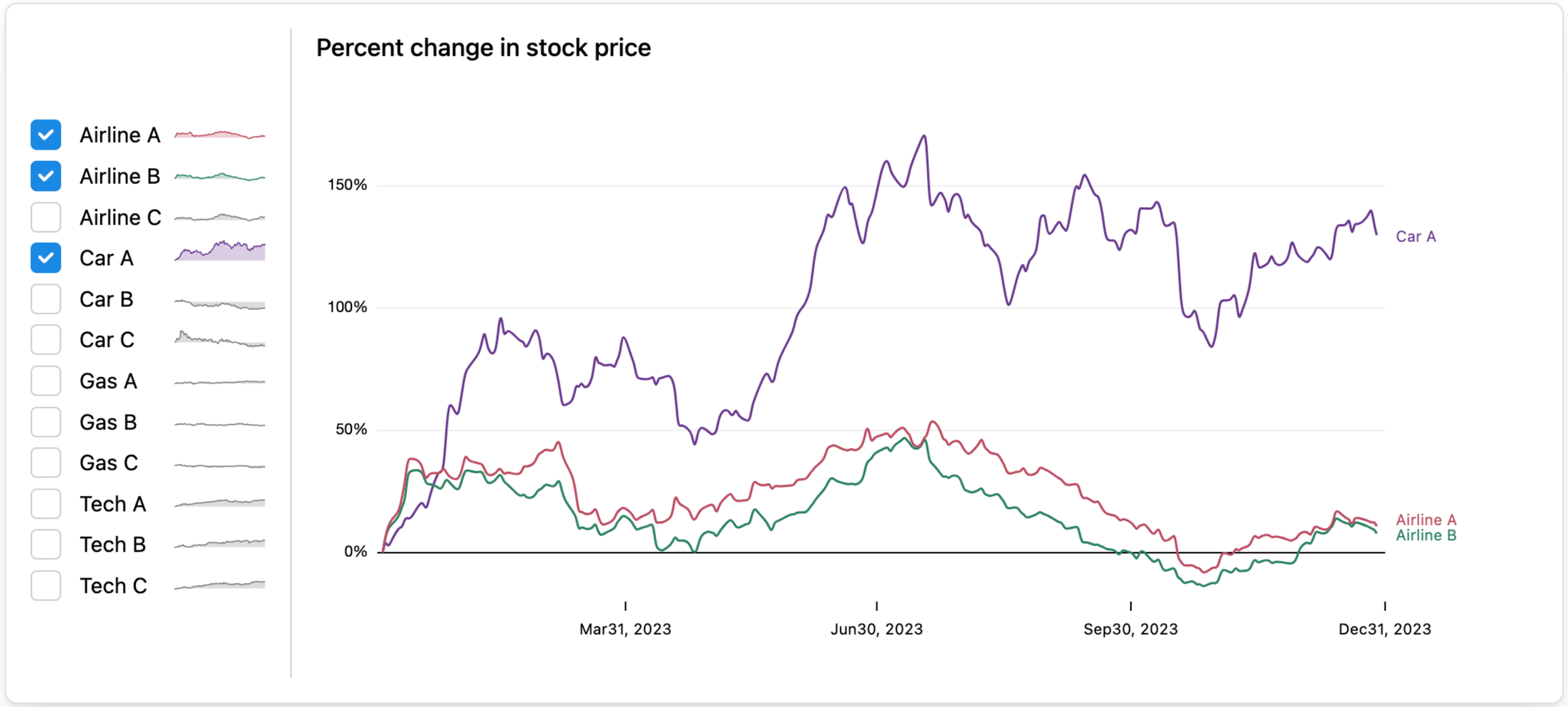
ENFORCING CONTEXT – SUPERIMPOSITION / PRIMARY DATA



SHOWING STATISTICAL CONTEXT — SUPERIMPOSITION / SUMMARY



CONTEXT OUTSIDE — JUXTAPOSITION / PRIMARY DATA



SUMMARY CONTEXT — JUXTAPOSITION / SUMMARY

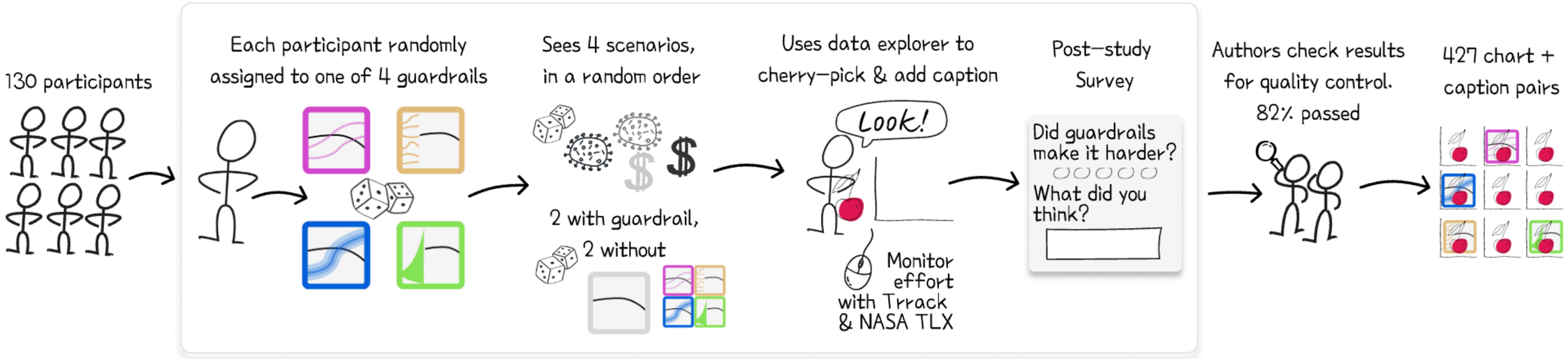
3 STUDIES

Producer Study: do guardrails make it more difficult to cherry-pick data?

Consumer Study: Do guardrails guardrails make cherry-picked data less convincing?

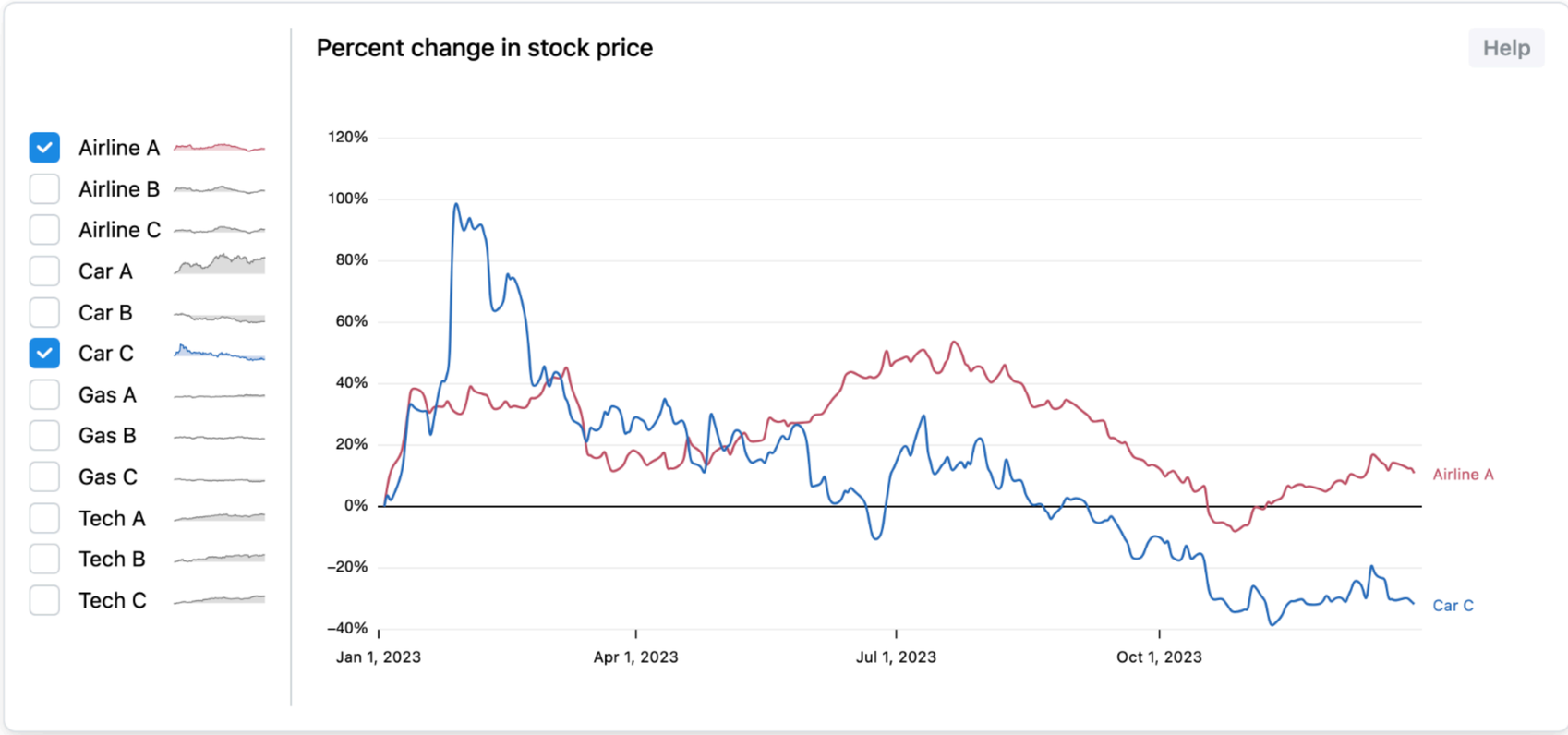
Controlled Consumer Study: Do guardrails work at different levels of egregiousness?

PRODUCER STUDY: SETUP



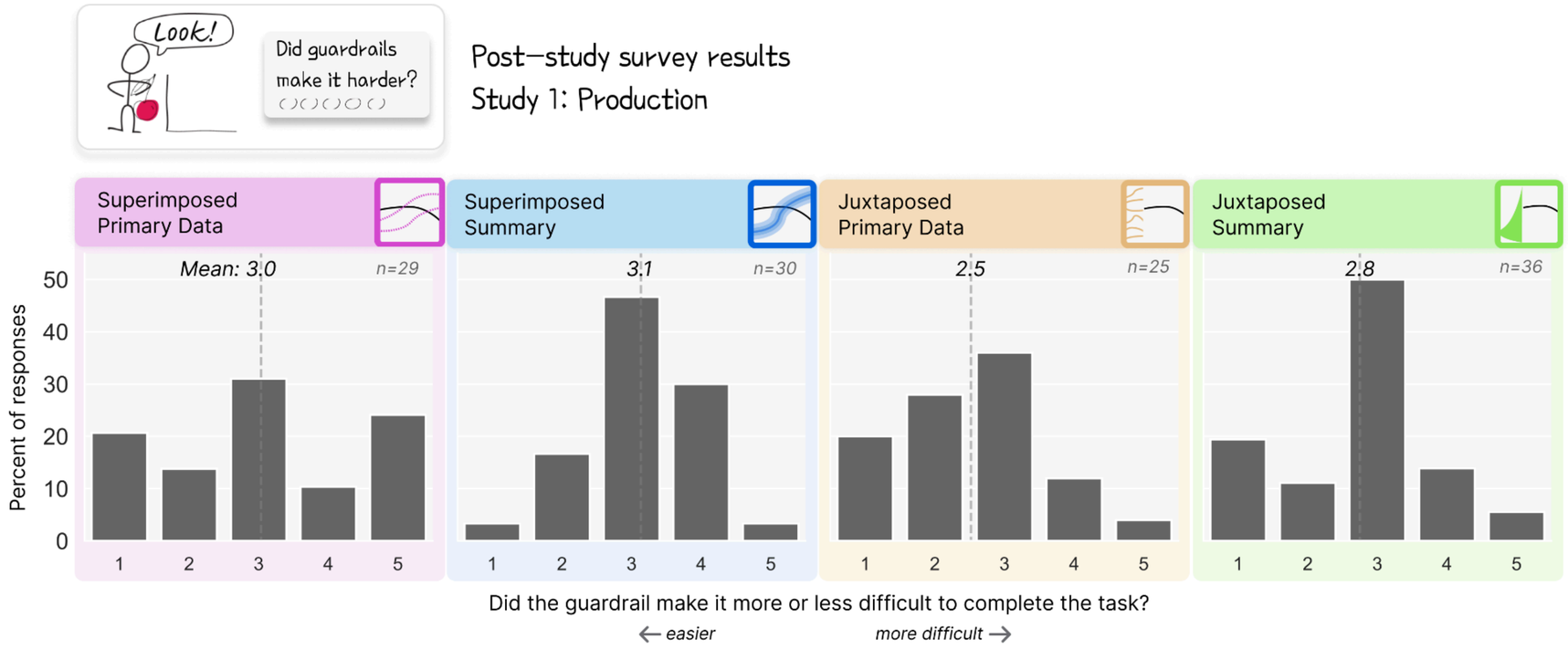
PRODUCER STUDY: SETUP

Select a view that best shows (and convinces your client) that **the airline industry fund is the best investment**.



Add a short text caption that will go along with your visualization: *

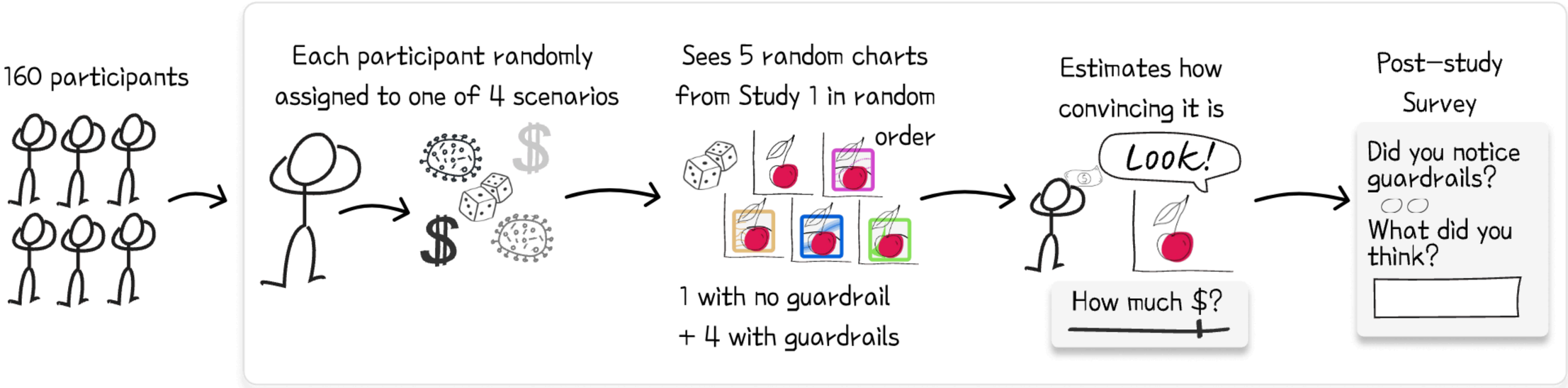
PRODUCER STUDY: RESULTS



Superimposed primary: **bimodal**

Juxtaposed primary: **makes cherry picking easier**

CONSUMER STUDY SETUP



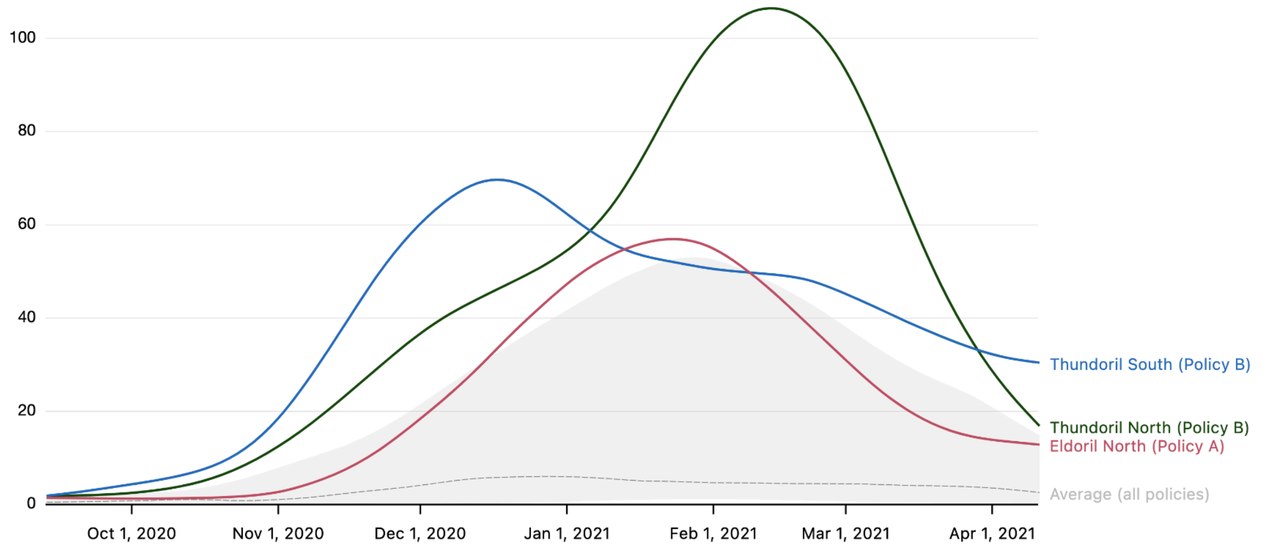
CONSUMER STUDY SETUP

> Policy A (red) is the superb policy to manage a sudden boom in infections

Infections per million people

Shaded area represents the middle 50% of all values.

Help



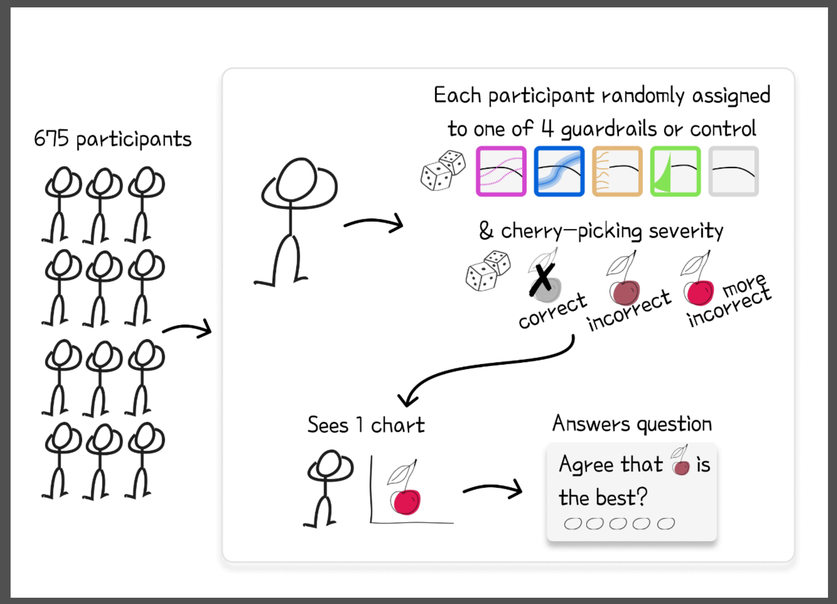
Based on this information, before traveling to **Eldoril North (Policy A)** I would buy this much insurance:
($0 = no risk of getting sick, $100 = very high risk of getting sick)

$0 $25 $50 $75 $100

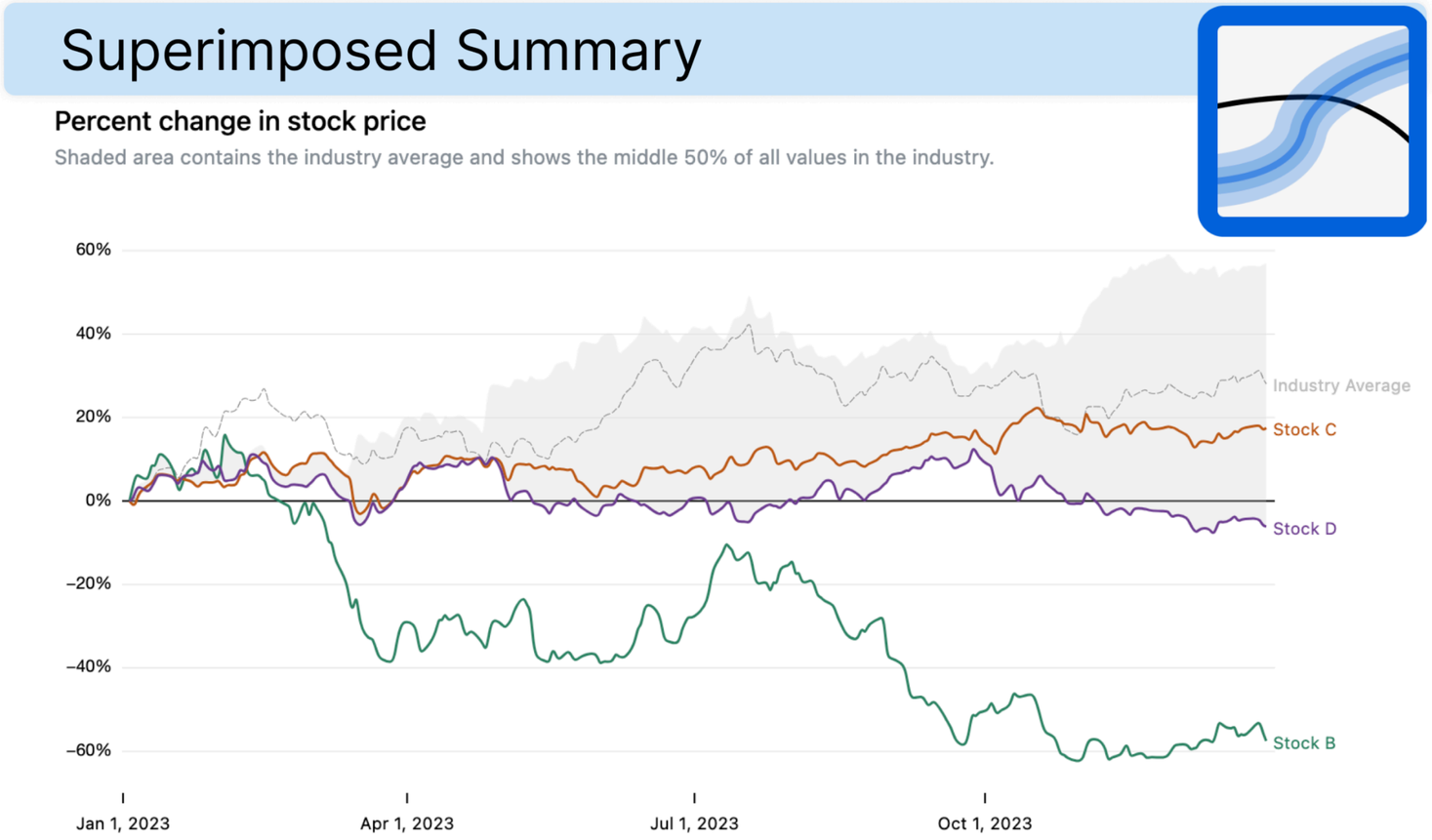
CONSUMER STUDY: RESULTS

Superimposed Primary works in some scenarios

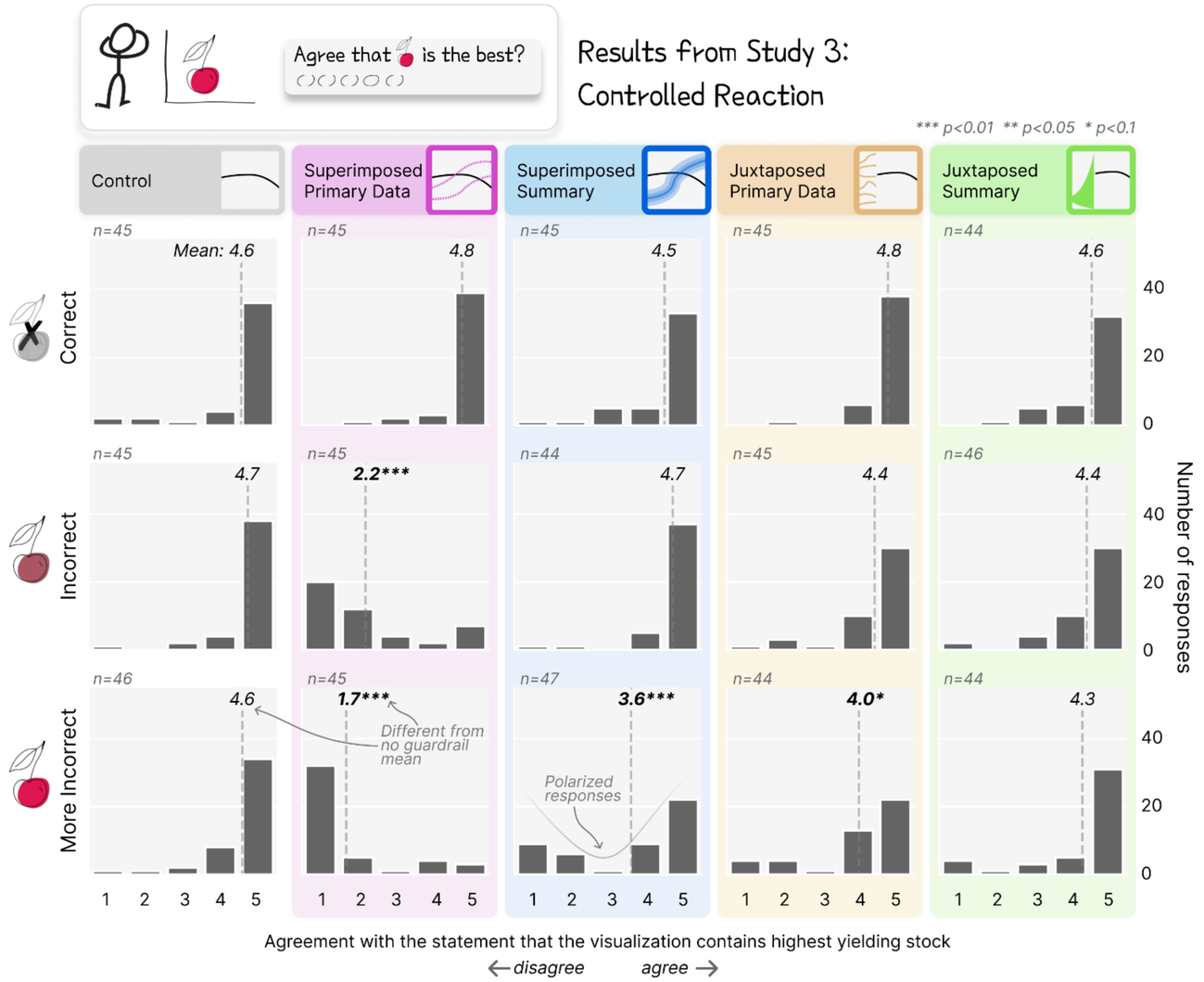
CONTROLLED CONSUMER STUDY



The visualization supports the idea that stock C yielded the highest returns



RESULTS



Superimposed primary works

Superimposed summary is bimodal

Juxtaposed doesn't work

DISCUSSION

Superimposed Primary seems like a promising design against cherry-picking

- easy to implement
- easy to understand

Juxtaposed Primary design useful as a way to make data explorers more usable

Interventions may be useful, but not a Panacea

SUMMARY

A TOOLKIT FOR DATA VISUALIZATION EXPERIMENTS

ReVISit takes care of the annoying parts of a study

Gives you the **ability to share your study with** reviewers and community!

Hope you give it a try.

Alexander Lex
http://vdl.sci.utah.edu

THE UNIVERSITY OF UTAH

The reVISit User Study Platform and Applications in Studying Misinformation

visualization design lab

SCI
www.sci.utah.edu