



Accelerating Oncology Data Review:

An R Shiny application for insights

Matthew Kumar
Srinivas Veeragoni

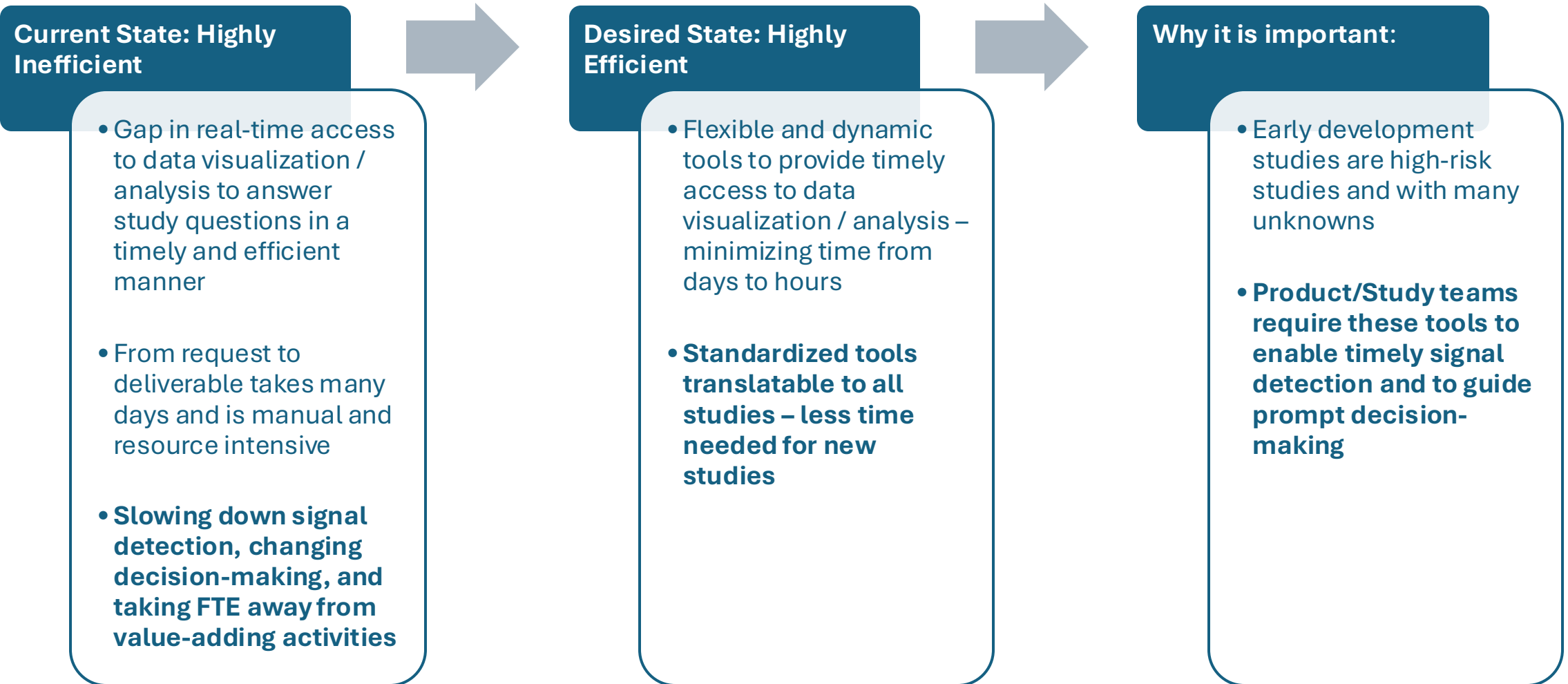
Sept 5, 2025





Fast access to clinical data analysis for efficient decision-making

to speed up development timelines





Data Navigator: Overview

Overview

- // An R Shiny application to identify and characterize clinical data for early phase oncology trials
- // **AIM:** Enable timely exploration of data, thereby *accelerating* decision making
- // **FOCUS 1:** Baseline data for patient population stratification
- // **FOCUS 2:** Evaluating treatment efficacy
- // Developed through iteratively and in collaboration with SMEs
- // Under active development and in MVP status. It is currently not validated (more later)
- // **IMPACT:** Delivers previously unavailable functionality to end users that accelerates review and exploration

Landing Page

Data Navigator

Study Information | Bar Chart | Box Chart | Waterfall Chart | Spider Chart | Swimmer Chart | Tornado Chart | AI Assistant

Study Information

Data Navigator

Welcome to the Landing Page

Your interactive application for exploring early baseline and efficacy data.
Formerly known as Efficacy Navigator.

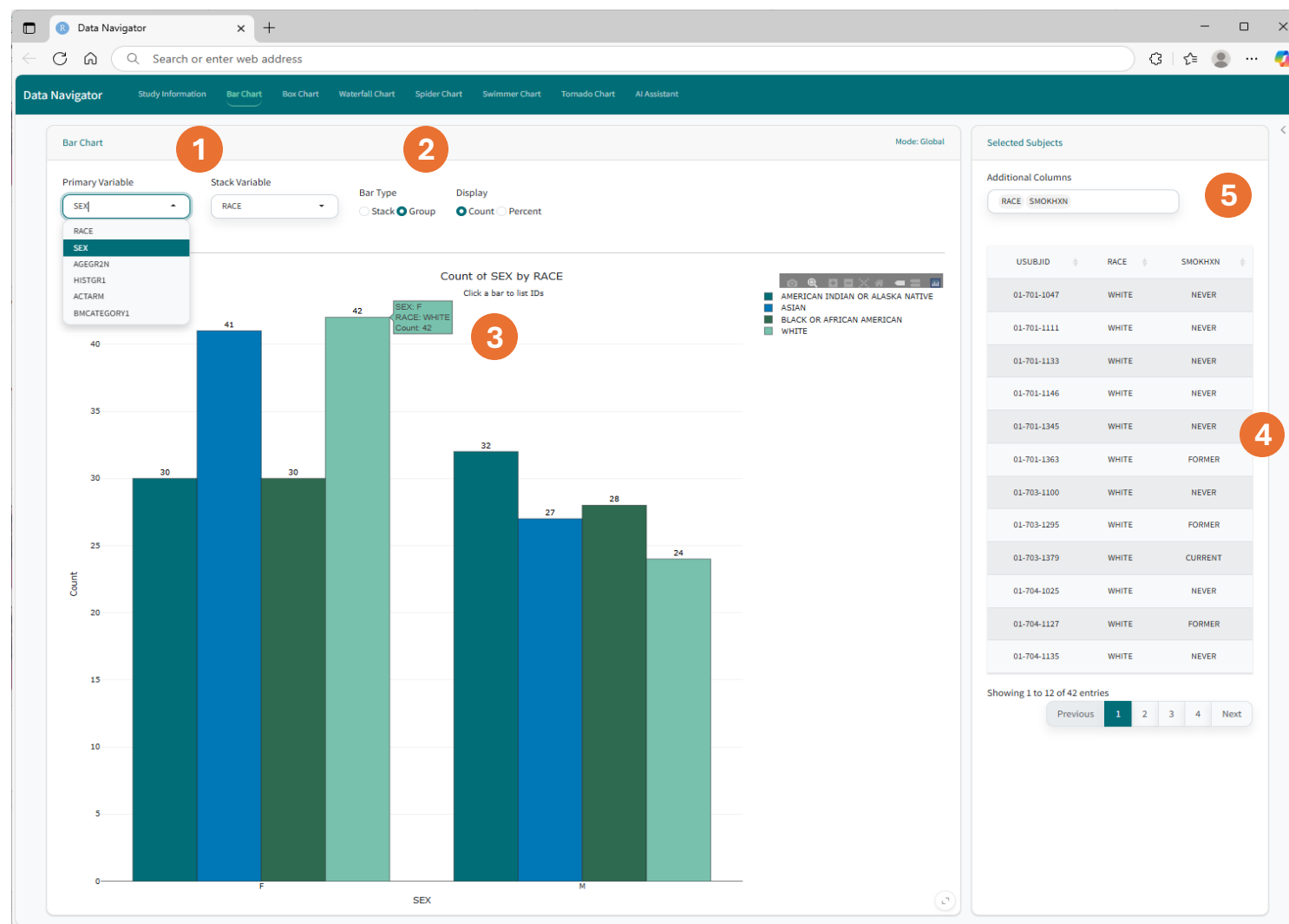
Study Number	99999
Bay Number	1234567
Project Description	A Mock Compound
Study Description	A Mock Study
Phase	1
Safety Analysis Set	Y: 254 (100%) / N: 0 (0%) / Missing: 0 (0%)
Efficacy Analysis Set	Y: 234 (92.1%) / N: 20 (7.9%) / Missing: 0 (0%)

[Set Global Filters](#)



Example Visualization Module: Bar Charts

Tool View



Steps

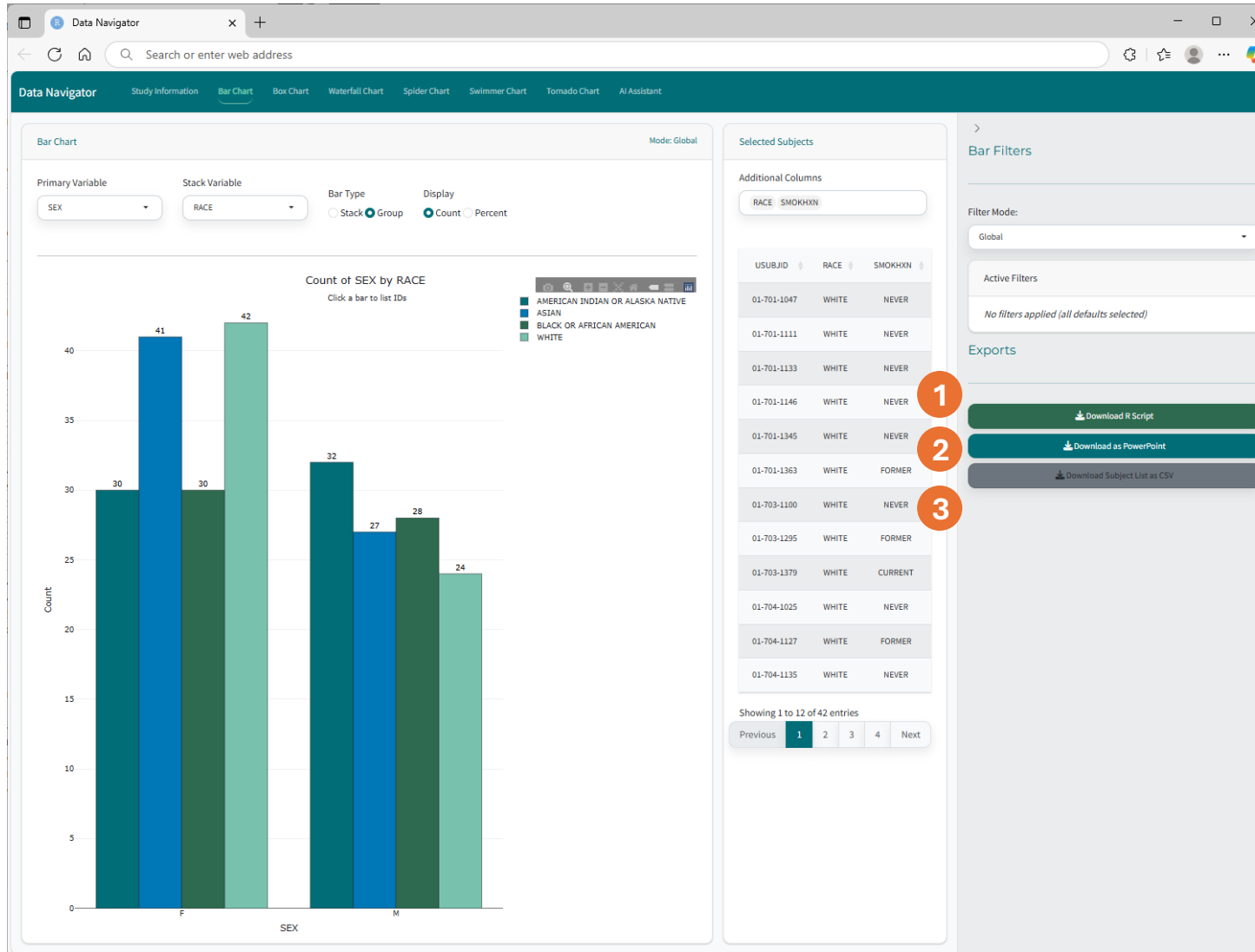
- 1 Select category variables for bar chart
- 2 Select bar chart display options
- 3 Tool tips provide contextual feedback
- 4 Clicking a bar reveals a table of subjects in that bar
- 5 Additional variables can be supplemented

Note: Data are from [Phuse TestDataFactory](#)



Example Visualization Module: Bar Charts - Exporting

Tool View



Steps

1. Export R Script that created *this* visual
2. Export to image and table to PPTX
3. Export Subject List to CSV





Example Visualization Module: Bar Charts - Output

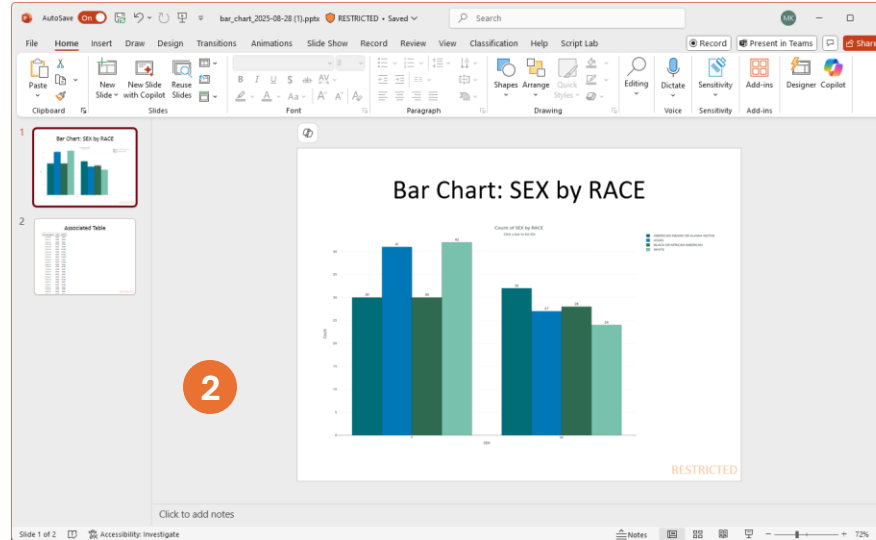
Output

```

1 # Bar Chart: SEX by RACE
2 # 2025-08-28 17:03:44.165658
3 # primary_var : SEX
4 # stack_var : RACE
5 # bar_mode : group
6 # y_scale : count
7 # Number of subjects: 254
8
9 # NOTE: Make sure 'demo_df' is loaded in your environment
10 # Example: demo_df <- readRDS('path/to/your/data.rds')
11
12 # Bar Chart Template - Complete R Script
13 # Required libraries
14 library(plotly)
15 library(dplyr)
16
17 # Filter to relevant subjects and select required columns
18 data <- demo_df %>%
19   filter(USUBJID %in% c(
20     "01-701-1015",
21     "01-701-1023",
22     "01-701-1028",
23     "01-701-1033",
24     "01-701-1034",
25     "01-701-1047",
26     "01-701-1097",
27     "01-701-1111",
28     "01-701-1115",
29     "01-701-1118",
30     "01-701-1130",
31     "01-701-1133",
32     "01-701-1146",
33     "01-701-1148",
34     "01-701-1153",
35     "01-701-1180",
36   )) %>%
37   select(SEX, RACE)
38
39 # Data preparation
40 chart_data <- demo_df %>%
41   count(.data[["SEX"]], .data[["RACE"]], name = "count") %>%
42   mutate(primary_cat = .data[["SEX"]],
43          stack_cat = .data[["RACE"]]) %>%
44   group_by(primary_cat) %>%
45   mutate(percent = count / sum(count) * 100) %>%
46   ungroup()

```

1



3

Steps

- 1 Export R Script that created *this* visual
- 2 Export to image and table to PPTX
- 3 Export Subject List to CSV

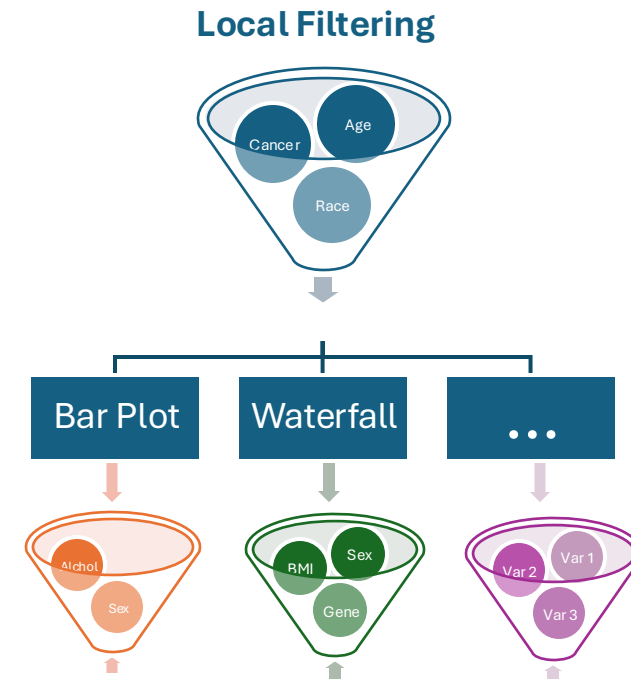
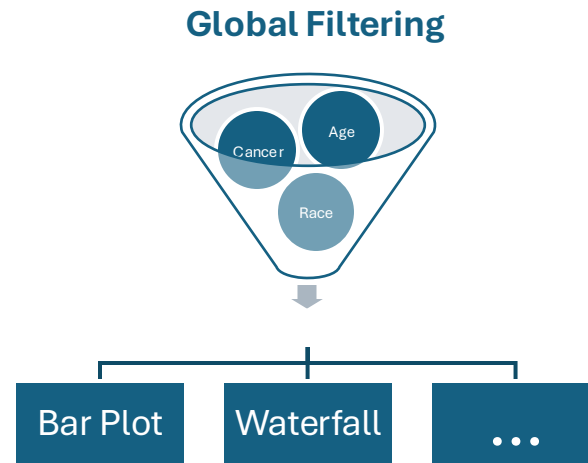




Data Navigator: Interactive Filtering

Filtering

- // Requirement that user can interactively segment data
- // **IMPACT:** Drives exploration, towards self-service analytics
- // A dual filtering strategy was implemented:



- // Tracking of active filters is imperative
- // Study-specific filters (e.g., extensibility) also needed to be accounted for



Data Navigator: Example Global Filter Setup

Tool View

The screenshot shows the Data Navigator tool interface. The main area displays a welcome message and a table with study information. A 'Set Global Filters' button is at the bottom. On the right, a 'Global Filters' panel is open, showing active filters and variable selection options. Red numbered circles 1, 2, and 3 highlight key features: 1 points to the Global Filters panel, 2 points to the 'Active Filters' section, and 3 points to the 'General Variables' section.

Study Information

Study Number	99999
Bay Number	1234567
Project Description	A Mock Compound
Study Description	A Mock Study
Phase	1
Safety Analysis Set	Y: 134 (100%) / N: 0 (0%) / Missing: 0 (0%)
Efficacy Analysis Set	Y: 124 (92.5%) / N: 10 (7.5%) / Missing: 0 (0%)

Global Filters

Active Filters

- Race: ASIAN, WHITE

General Variables

- Subject ID: 254 items selected
- Age: 51 to 89
- Pooled Age Group 2: <65, ≥65
- Sex: F, M
- Race: ASIAN, WHITE (selected), AMERICAN INDIAN OR ALASKA NATIVE, BLACK OR AFRICAN AMERICAN, WHITE

TAS Variables

- Efficacy Analysis Set: N, Y
- Safety Analysis Set: Y
- Histology Group: (empty)

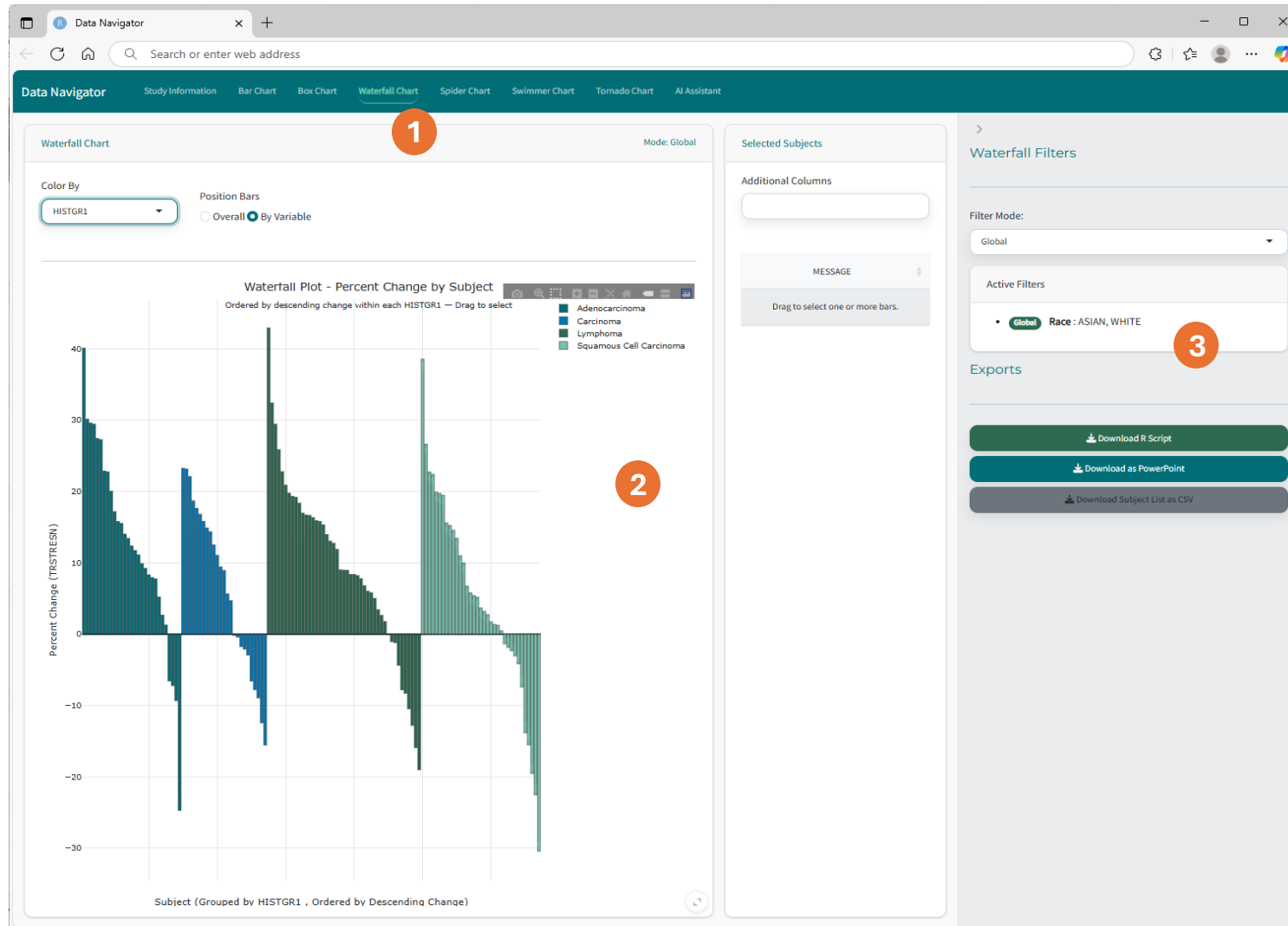
Steps

- 1 User opens Global Filter panel from landing page
- 2 User freely chooses values to filter on (multiple are possible)
- 3 A status indicator lists all active filters



Data Navigator: Using Filtered Data

Tool View



Steps

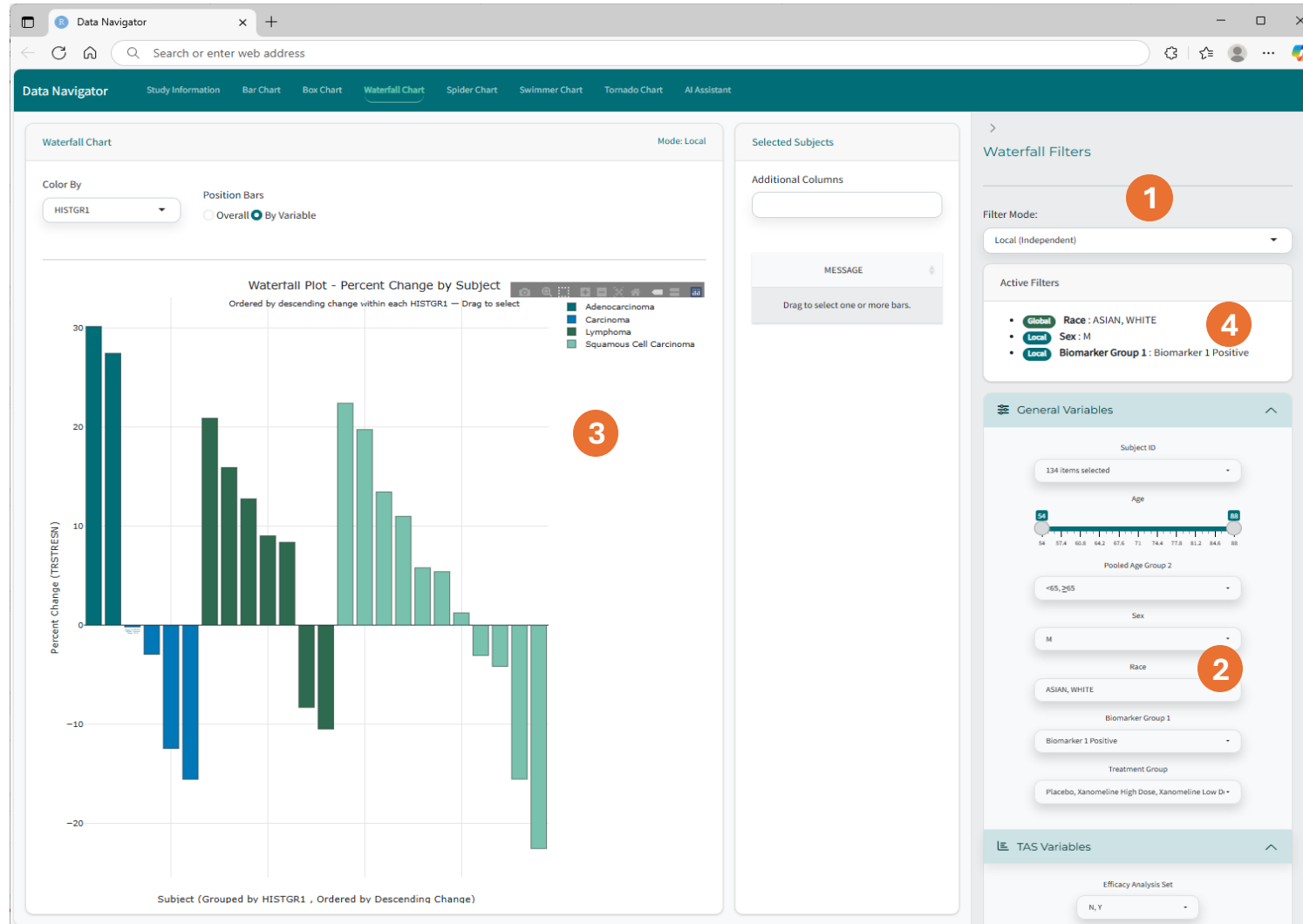
- 1 User navigates to Waterfall Chart
- 2 Global filter already applied to visualization
- 3 Status indicator reminds user what filters are active





Data Navigator: Locally Filtered Data

Tool View



Steps

- 1 User switches to Local filtering within Waterfall Chart side panel
- 2 An identical set of filters is revealed; user may select *additional* filters
- 3 Waterfall Chart updated with *additional* filters
- 4 Status indicator tracks all active filters on the Waterfall Chart



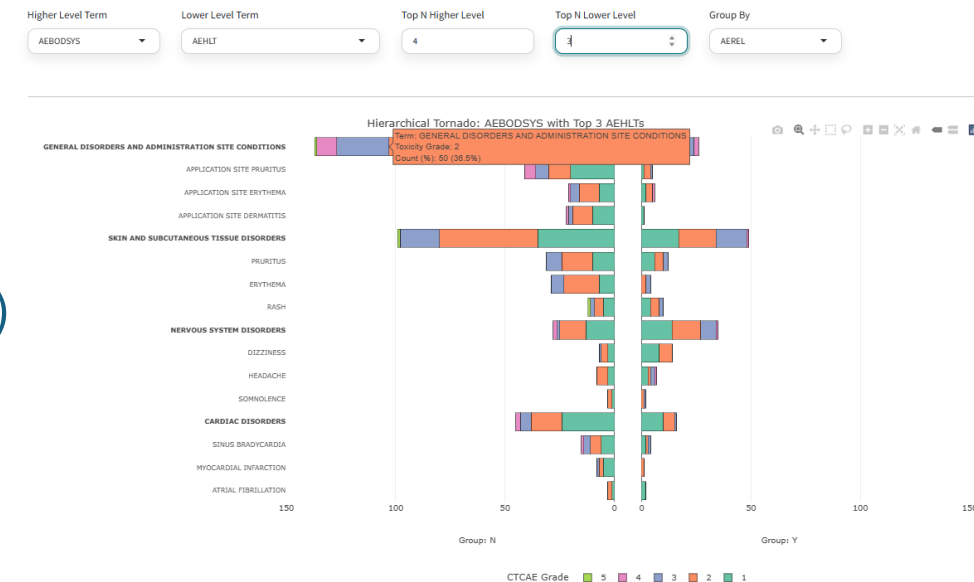
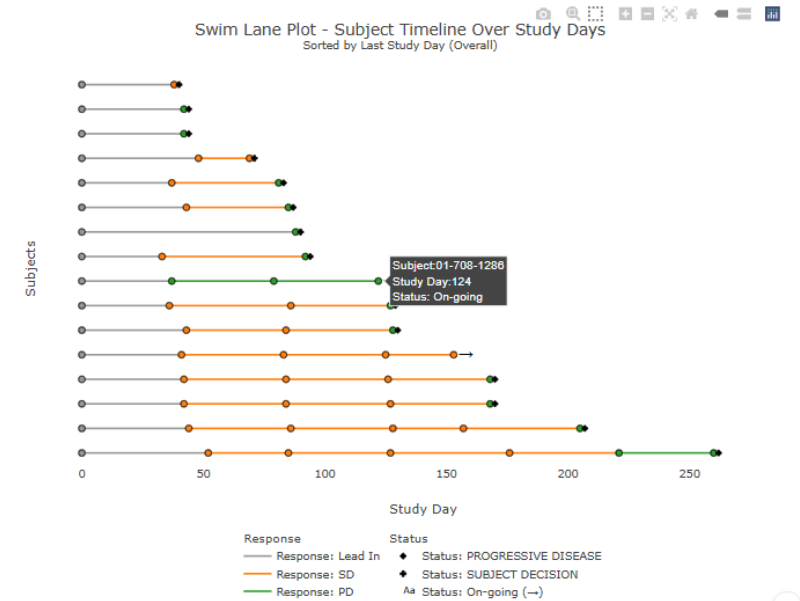
Note: All exports work with applied filters



Other Modules

Overview

- // The modular approach employed makes extending base functionalities to other visualizations faster
- // Ensures a clear and consistent UI and UX that delivers timely results
- // The DX also benefits ➔ new modules are easier to bootstrap, test, document
- // The overall base application is driven by a YAML configuration file
- // Provisions for filters specification (default variables + values and study specific variables)
- // Ensures scalability to the next study (data model is fixed)





What If?

Overview

- // Filtering is quite extensive already, but there are cases where it can be tricky (e.g., filter subjects above the average age)
- // Each visualization is designed with a specific focus. While interactive, they are still “fixed” in a sense (e.g., by design)
- // **TAKEAWAY:** There will always be “what if” and one-off questions that are *not* capable of being answered on the spot

- // **IDEA:** Can we leverage AI for these one-off tasks?
- // In short, yes. But much consideration and work needs to be done to bring this to a truly safe and productive use!
- // The next slides show how we’ve integrated an **AI Assistant** to help with novel:
 - // Question + Answering
 - // Data Analysis
 - // Data Visualization
- // This is experimental



Data Navigator: AI Assistant Space

Tool View

Data Navigator

Study Information | Bar Chart | Box Chart | Waterfall Chart | Spider Chart | Swimmer Chart | Tornado Chart | **AI Assistant**

Data Space

USUBJID	AGE	SEX	RACE	ETHNIC	ACTARM	AGEGR2N	BMCATEGORY1	HEIGHTBL
01-701-1015	63	F	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO	Placebo	<65	Biomarker 1 Negative	58
01-701-1023	64	M	AMERICAN INDIAN OR ALASKA NATIVE	NOT HISPANIC OR LATINO	Placebo	<65	Biomarker 1 Positive	64
01-701-1028	71	M	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO	Xanomeline High Dose	≥65	Biomarker 1 Positive	70
01-701-1033	74	M	AMERICAN INDIAN OR ALASKA NATIVE	NOT HISPANIC OR LATINO	Xanomeline Low Dose	≥65	Biomarker 1 Negative	69
01-701-1034	77	F	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO	Xanomeline High Dose	≥65	Biomarker 1 Positive	61

Plotting Space

No plot available
Ask the AI to create a visualization!

AI Assistant

Hello! 🤖 I'm here to help you analyze and explore the clinical trial data in this dashboard. I can assist you with several types of tasks:

Filtering:

- Filter to show females with Adenocarcinoma +
- Filter on subjects above the overall average age +

★ To remove filters from the dashboard, simply say reset filters ★

Data Analysis Questions:

- How many histology groups are present +
- What is the average baseline bmi by age and sex +

Visualizations:

- Create a histogram of patient weight, facet by ECOG +
- Plot age versus weight colored by treatment group as a scatter plot +

What would you like to explore first?

Enter a message...

Steps

- 1 User is presented a chat interface to interact with AI Assistant
- 2 Data Space contains an interactive table, reacts to filter queries
- 3 Plotting Space contains a blank canvas, reacts to visualization queries





Data Navigator: Filtering in AI Assistant Space

Tool View

The screenshot shows the Data Navigator interface. On the left, the 'Data Space' section displays a table of clinical trial data. The table has columns: USUBJID, AGE, SEX, RACE, ETHNIC, ACTARM, AGEGR2N, BMCATEGORY1, and HEIGHTBL. The data is filtered to show only white males in the placebo arm. A red box highlights the filtered data, and a red circle with the number 3 is placed over the table header.

USUBJID	AGE	SEX	RACE	ETHNIC	ACTARM	AGEGR2N	BMCATEGORY1	HEIGHTBL
01-701-1234	69	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	71.5
01-708-1253	61	M	WHITE	NOT HISPANIC OR LATINO	Placebo	<65	Biomarker 1 Positive	73
01-710-1027	83	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	67
01-710-1264	78	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	66
01-711-1036	70	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	67.7007874015748
01-716-1177	72	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	70
01-718-1139	77	M	WHITE	NOT HISPANIC OR LATINO	Placebo	≥65	Biomarker 1 Positive	67.5

Below the table is the 'Plotting Space' section, which currently shows 'No plot available. Ask the AI to create a visualization!'.

On the right, the 'AI Assistant' chat interface is shown. It contains a welcome message and a list of tasks. The first task is 'Filtering', which includes a filter query for 'white males in placebo arm'. The second task is 'Data Analysis Questions', which includes questions about histology groups and baseline BMI. The third task is 'Visualizations', which includes a histogram of patient weight and a scatter plot of age versus weight. The chat interface also shows a response to the filter query, including the SQL code and a confirmation message.

Filtering:

- Filter to show females with Adenocarcinoma
- Filter on subjects above the overall average age

★ To remove filters from the dashboard, simply say [reset filters](#) ★

Data Analysis Questions:

- How many histology groups are present
- What is the average baseline bmi by age and sex

Visualizations:

- Create a histogram of patient weight, facet by ECOG
- Plot age versus weight colored by treatment group as a scatter plot

What would you like to explore first?

1 Filter to show white males in the placebo arm

2

```
SELECT * FROM data_for_11m
WHERE RACE = 'WHITE' AND SEX = 'M' AND ACTARM = 'Placebo'
```

The dashboard has been filtered to show only white males in the placebo arm.

Enter a message...

Steps

- 1 Filter query for “white males in placebo arm” submitted
- 2 Response includes SQL code
- 3 Data Space table is automatically filtered





Data Navigator: Filtering in AI Assistant Space

Tool View

The screenshot shows the Data Navigator interface. On the left, there is a 'Data Space' section with a table of patient data. The 'AGE' column is highlighted with a red box. Below the table is a 'Plotting Space' section with a message: 'No plot available. Ask the AI to create a visualization!'. On the right, there is an 'AI Assistant' chat interface. The chat shows a conversation where the user asks to filter subjects above the average age, and the assistant responds with the SQL code and the filtered data.

USUBJID	AGE	SEX	RACE	ETHNIC	ACTARM	AGEGRN	BMCATEGORY1	HEIGHTBL	WEIGHTBL	BMIBL	ECOGBL
01-701-1034	77	F	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO	Xanomeline High Dose	265	Biomarker 1 Positive	61	138	26.0720236495566	0
01-701-1047	85	F	WHITE	NOT HISPANIC OR LATINO	Placebo	265	Biomarker 1 Positive	58.5	148	30.4022207611951	0
01-701-1111	81	F	WHITE	NOT HISPANIC OR LATINO	Xanomeline Low Dose	265	Biomarker 1 Negative	62.2992125984252	132	23.90915022355026	0
01-701-1115	84	M	AMERICAN INDIAN OR ALASKA NATIVE	NOT HISPANIC OR LATINO	Xanomeline Low Dose	265	Biomarker 1 Negative	71.5	174	23.927233605553	0
01-701-1130	84	M	ASIAN	NOT HISPANIC OR LATINO	Placebo	265	Biomarker 1 Negative	66.7992125984252	175	27.5708837608556	0
01-701-1133	81	F	WHITE	NOT HISPANIC OR LATINO	Xanomeline High Dose	265	Biomarker 1 Negative	64	164	28.1474609375	0
01-701-1153	79	F	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO	Placebo	265	Biomarker 1 Positive	61.5	128	23.7911296186133	0

AI Assistant Chat:

I can assist you with several types of tasks:

Filtering:

- Filter to show females with Adenocarcinoma
- Filter on subjects above the overall average age

★ To remove filters from the dashboard, simply say [reset filters](#) ★

Data Analysis Questions:

- How many histology groups are present
- What is the average baseline bmi by age and sex

Visualizations:

- Create a histogram of patient weight, facet by ECOG
- Plot age versus weight colored by treatment group as a scatter plot

What would you like to explore first?

1 filter on subjects above of the overall age. inform me of the value first

`SELECT AVG(AGE) AS average_age FROM data_for_11m`

average_age
75.09

The average age of the subjects is approximately 75.09 years. Would you like me to filter the data to show only subjects above this age?

yes

2

`SELECT * FROM data_for_11m WHERE AGE > 75.0866`

I've filtered the dashboard to show only subjects who are above the average age of 75.09 years.

Enter a message...

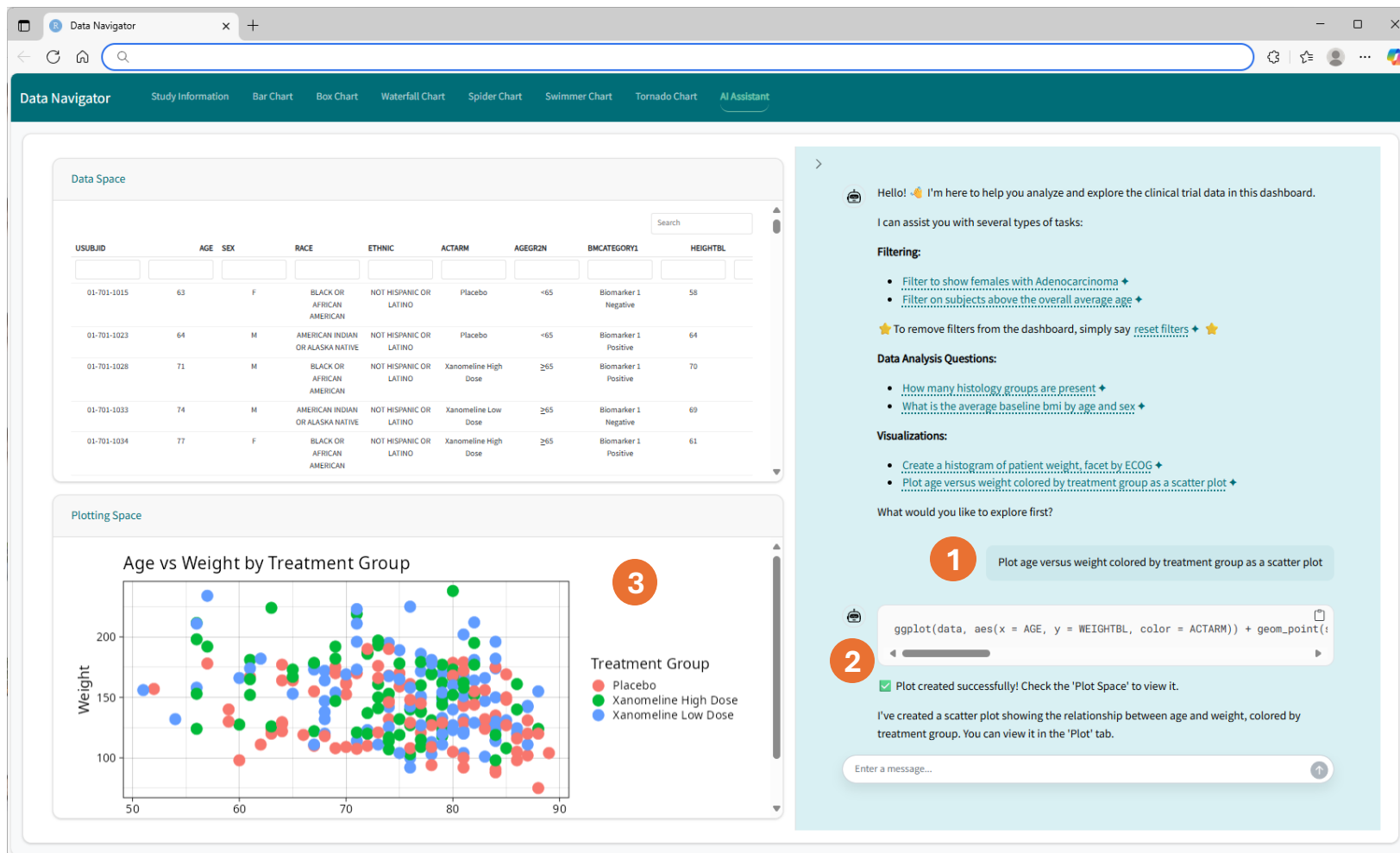
Steps

- 1 Filter query for “subjects above the average age” submitted
- 2 Response includes SQL code, and “verification” step
- 3 Data Space table is automatically filtered



Data Navigator: Visualization in AI Assistant Space

Tool View



Steps

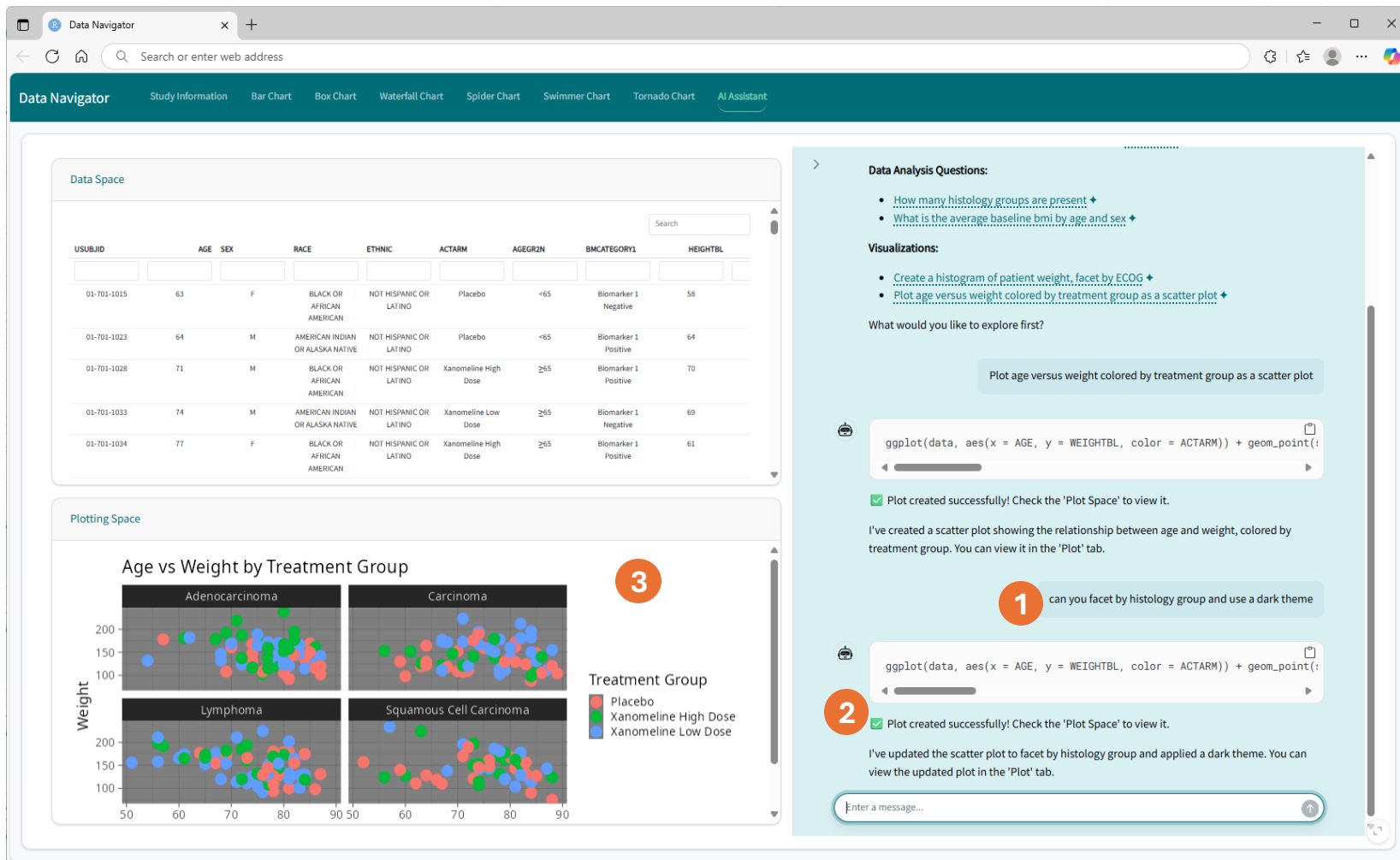
- 1 Visualization query for "scatterplot" submitted
- 2 Response includes R/ggplot2 code
- 3 Plotting Space canvas renders the visualization





Data Navigator: Visualization in AI Assistant Space

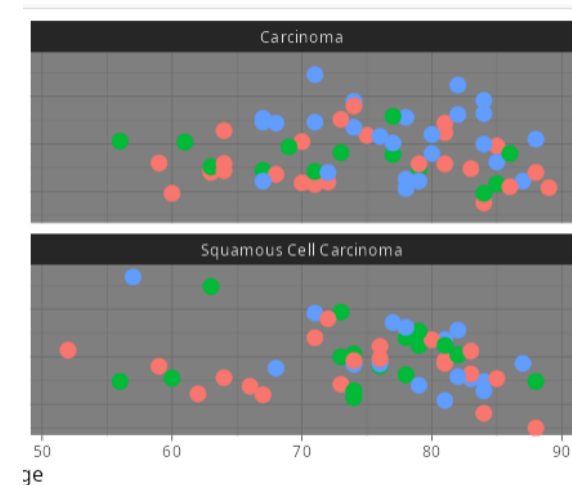
Tool View



Steps

- 1 Follow up visualization query for "scatterplot" submitted
- 2 Response includes *updated* R/ggplot2 code
- 3 Plotting Space produces the *updated* visual

Note: All images have a watermark





Ecosystem

Brief Overview

Environment, Data Engineering, Data Model



databricks



posit

Application Packages



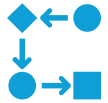
Internal Platforms, Services





Next Steps

Road Map



Continue to iterate on requested visualization set, functionality



Roll out application to 3 pilot studies



Migration to GxP-environment scheduled



Incorporate safety data and analysis requirements



Integration of additional data sources (RWE, 2ndary use, digital biomarker)



Continue to explore the role of AI with agentic approaches



Thank You

Acknowledgements:

Early Clinical Development Oncology (ECDO) Team
DSE TITANS Team
Data Monitoring and Operational Insights (DMOI) Team
Statistical Programing & Analytics (SPA) and Statistics Team

