

# *Graph Designer: A R shiny application to create publication quality graphs in minutes*

***Mukul Mittal, PhD***

*mukul.mittal@sanofi.com*

# Background:

---

## What

Graph designer is a R Shiny-based application.

---

Application adopts “**The Grammar of Graphics**” principle using **ggplot2** package for graph creation.

---

Industry first ggplot2 based dynamic code re-generative visualization tool.

---

## Why

To streamline, accelerate and equip biostatisticians with a tool to create standard publication ready visualizations within minutes vs hours.

---

To provide a powerful, time and code-efficient visualization tool to aid in studies across business units, verticals, and regions.

---

Self-learning tool to study, understand and learn the in-and-outs of visualization using R {ggplot2} & create simple-to-complex viz for analysis.

---

# R Shiny:

- Shiny is an R package, enabling data scientists to build interactive web applications to execute R code at the backend.
- Easy to integrate custom UI/UX web development components via HTML/CSS/JS.
- Free, open-source, and widely-used web development framework in R.
- Continuous Support and development from RStudio catering to bugs, issues and fixes.
- Flexibility to setup, either local or on remote-server, and deploy with custom IT solutions.
- Plethora of open-source supporting packages to introduce, improve and build add-ins for front-end and back-end functionality.
- Active public R and Shiny community.
- #1 framework for rapid development by pharma organizations across the globe.

# Application Layout:

 **Graph Designer**

1 Upload Data

2 Review Your Data

3 Data Filter

4 Visualize

5 Publish

Powered by  **sanofi**

Feedback    

## How do you want to upload your data?

Using the tab set below, **select a read method** and **upload your data** into the application. The uploaded data will be minimally populated in the editable text box to the bottom right.

Once uploaded, you will be able to **view your data** and, if needed, interactively **define column data types** (e.g. numeric, character).

Note, if data has already been uploaded, you can use the same tab set to “update” your data (assuming the columns names stay the same).

Demo  Copy & Paste  CSV Upload  Excel Upload  RDS Upload  SAS Upload  XPT Upload

### Demo Examples

If you just want to try Graph Designer, select from the list of example datasets on the bottom.

Select a demo dataset

adsl

Verify/Modify your data:

Go To Review Your Data

|    |                                                                                                                        |
|----|------------------------------------------------------------------------------------------------------------------------|
| 1  | "STUDYID", "USUBJID", "SUBJID", "SITEID", "SITEGR1", "ARM", "TRT01P", "TRT01PN", "TRT01A", "TRT01AN", "TRTSDT", "TRTE  |
| 2  | "CDISCPIL0T01", "01-701-1015", "1015", "701", "701", "Placebo", "Placebo", 0, "Placebo", 0, 2014-01-02, 2014-07-02, 18 |
| 3  | "CDISCPIL0T01", "01-701-1023", "1023", "701", "701", "Placebo", "Placebo", 0, "Placebo", 0, 2012-08-05, 2012-09-01, 28 |
| 4  | "CDISCPIL0T01", "01-701-1028", "1028", "701", "701", "Xanomeline High Dose", "Xanomeline High Dose", 81, "Xanomeli     |
| 5  | "CDISCPIL0T01", "01-701-1033", "1033", "701", "701", "Xanomeline Low Dose", "Xanomeline Low Dose", 54, "Xanomeline     |
| 6  | "CDISCPIL0T01", "01-701-1034", "1034", "701", "701", "Xanomeline High Dose", "Xanomeline High Dose", 81, "Xanomeli     |
| 7  | "CDISCPIL0T01", "01-701-1047", "1047", "701", "701", "Placebo", "Placebo", 0, "Placebo", 0, 2013-02-12, 2013-03-09, 26 |
| 8  | "CDISCPIL0T01", "01-701-1097", "1097", "701", "701", "Xanomeline Low Dose", "Xanomeline Low Dose", 54, "Xanomeline     |
| 9  | "CDISCPIL0T01", "01-701-1111", "1111", "701", "701", "Xanomeline Low Dose", "Xanomeline Low Dose", 54, "Xanomeline     |
| 10 | "CDISCPIL0T01", "01-701-1115", "1115", "701", "701", "Xanomeline Low Dose", "Xanomeline Low Dose", 54, "Xanomeline     |
| 11 | "CDISCPIL0T01", "01-701-1118", "1118", "701", "701", "Placebo", "Placebo", 0, "Placebo", 0, 2014-03-12, 2014-09-09, 18 |
| 12 | "CDISCPIL0T01", "01-701-1130", "1130", "701", "701", "Placebo", "Placebo", 0, "Placebo", 0, 2014-02-15, 2014-08-16, 18 |
| 13 | "CDISCPIL0T01", "01-701-1133", "1133", "701", "701", "Xanomeline High Dose", "Xanomeline High Dose", 81, "Xanomeli     |
| 14 | "CDISCPIL0T01", "01-701-1146", "1146", "701", "701", "Xanomeline High Dose", "Xanomeline High Dose", 81, "Xanomeli     |
| 15 | "CDISCPIL0T01", "01-701-1148", "1148", "701", "701", "Xanomeline High Dose", "Xanomeline High Dose", 81, "Xanomeli     |

**sanofi**

## Review your data:

Graph Designer Powered by  sanofi

1 Upload Data 2 Review Your Data 3 Data Filter 4 Visualize 5 Publish Feedback ⓘ 🔍 ?

Review your data...

Now that your data is uploaded, please review your data.

In the **View** section, the data is displayed in an interactive web table. In the **Column Types** section, if necessary, coerce the column types (e.g. numeric -> character), which will have a direct impact in the graph design process. In the **Structure** and **Summary** sections, review the structure and diagnostics of the data.

Once your review is completed, initialize the visualization by clicking the following button:

View

Show 5 entries

| STUDYID       | USUBJID     | SUBJID | SITEID | SITEGR1 | ARM                  | TRT01P               | TRT01PN | TRT01A               | TRT01AN | TRTSDT     | TRTEDT     | TRTDUR | AVGDD                                                                                                                                                                            | CUMDOSE | AGE | AGEGR1 | AGEGR1N | AGEU  | RACE  |
|---------------|-------------|--------|--------|---------|----------------------|----------------------|---------|----------------------|---------|------------|------------|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|-----|--------|---------|-------|-------|
| CDISCIPILOT01 | 01-701-1015 | 1015   | 701    | 701     | Placebo              | Placebo              | 0       | Placebo              | 0       | 2014-01-02 | 2014-07-02 | 182    | 0.0                                                                                                                                                                              | 0       | 63  | <65    | 1       | YEARS | WHITE |
| CDISCIPILOT01 | 01-701-1023 | 1023   | 701    | 701     | Placebo              | Placebo              | 0       | Placebo              | 0       | 2012-08-05 | 2012-09-01 | 28     | 0.0                                                                                                                                                                              | 0       | 64  | <65    | 1       | YEARS | WHITE |
| CDISCIPILOT01 | 01-701-1028 | 1028   | 701    | 701     | Xanomeline High Dose | Xanomeline High Dose | 81      | Xanomeline High Dose | 81      | 2013-07-19 | 2014-01-14 | 180    | <div><div></div><div>Graph Designer</div></div> <div><div>1 Upload Data</div><div>2 Review Your Data</div><div>3 Data Filter</div><div>4 Visualize</div><div>5 Print</div></div> |         |     |        |         |       |       |
| CDISCIPILOT01 | 01-701-1033 | 1033   | 701    | 701     | Xanomeline Low Dose  | Xanomeline Low Dose  | 54      | Xanomeline Low Dose  | 54      | 2014-03-18 | 2014-03-31 | 14     | <div><div>Column Types</div><div></div></div>                                                                                                                                    |         |     |        |         |       |       |


**Graph Designer**

Powered by  **sanofi**

1 Upload Data
2 Review Your Data
3 Data Filter
4 Visualize
5 Publish

[Feedback](#)
[Help](#)
[Info](#)
[Share](#)

### Column Types

|         |           |
|---------|-----------|
| STUDYID | character |
| USUBJID | character |
| SUBJID  | character |
| SITEID  | character |
| SITEGR1 | character |
| ARM     | character |
| TRT01P  | character |
| TRT01PN | integer   |
| TRT01A  | character |
| TRT01AN | integer   |
| TRTSOT  | Date      |
| TRTEDT  | Date      |

### Structure

```
'data.frame': 254 obs. of 48 variables:
 $ STUDYID : chr "CDISCPILOT01" "CDISCPILOT01" "CDISCPILOT01" ...
 $ USUBJID : chr "01-701-1015" "01-701-1023" "01-701-1028" "01-701-1033" ...
 $ SUBJID : chr "1015" "1023" "1028" "1033" ...
 $ SITEID : chr "701" "701" "701" "701" ...
 $ SITEGR1 : chr "701" "701" "701" "701" ...
 $ ARM : chr "Placebo" "Placebo" "Xanomeline High Dose" "Xanomeline Low Dose" ...
 $ TRT01P : chr "Placebo" "Placebo" "Xanomeline High Dose" "Xanomeline Low Dose" ...
 $ TRT01PN : int 0 0 81 54 81 0 54 54 54 0 ...
 $ TRT01A : chr "Placebo" "Placebo" "Xanomeline High Dose" "Xanomeline Low Dose" ...
 $ TRT01AN : int 0 0 81 54 81 0 54 54 54 0 ...
 $ TRTSOT : Date, format: "2014-01-02" "2012-08-05" ...
 $ TRTEDT : Date, format: "2014-07-02" "2012-09-01" ...
 $ TRTOUT : int 182 28 180 14 183 26 190 10 55 182 ...
 $ AVDOO : num 0 0 77.7 54 76.9 0 54 54 54 0 ...
 $ CUPDOSE : int 0 0 13986 756 14067 0 10269 540 2970 0 ...
 $ AGE : int 63 64 71 74 77 85 68 81 84 52 ...
 $ AGEGR1 : chr "<65" "<65" "65-80" "65-80" ...
 $ AGEGR1N : int 1 1 2 2 2 3 2 3 3 1 ...
 $ AGEU : chr "YEARS" "YEARS" "YEARS" "YEARS" ...
 $ RACE : chr "WHITE" "WHITE" "WHITE" "WHITE" ...
 $ RACEN : int 1 1 1 1 1 1 1 1 1 1 ...
 $ SEX : chr "M" "M" "M" "M" ...
 $ ETHNIC : chr "HISPANIC OR LATINO" "HISPANIC OR LATINO" "NOT HISPANIC OR LATINO" "NOT HISPANIC OR LATINO" ...
 $ SAFFL : chr "Y" "Y" "Y" "Y" ...
 $ ITTFL : chr "Y" "Y" "Y" "Y" ...
 $ EFFFL : chr "Y" "Y" "Y" "Y" ...
 $ COMBFL : chr "Y" "N" "Y" "N" ...
 $ COM16FL : chr "Y" "N" "Y" "N" ...
 $ COM24FL : chr "Y" "N" "Y" "N" ...
 $ DISCONF : chr "" "" "" "" ...
 $ DSDREFL : chr "" "" "" "" ...
 $ DTHFL : chr "" "" "" "" ...
 $ BMI : num 25.1 30.4 31.4 28.8 26.1 30.4 27.3 23.9 23.9 21.9 ...
```

# Data filtration & sorting:

## Filter your data...

Go To Visualization

Using this tab, you are able to filter your data so that you can choose a smaller part of your data set and use that subset for viewing, plotting or analysis.

Information related to filtered data is shown in the **summary** section. While filtering the data make sure that in your dataset there **should be no columns present which contains only null values otherwise you will get nothing in your data view because null value can't be filtered**.

While sorting your data in **select column to sort** field, it will be recommended that you should always use character or categorical variable. **The choose order to sort** field will allow you to apply some custom ordering of variable which can be visible in your graph.

**Note: The ascending or descending section in custom ordering will work only for selected variable in the field**

Once your review is completed, initialize the visualization by clicking the following button:

SELECT VARIABLES

STUDYID, USUBJID, SUBJID, SITEID, SI

STUDYID

NA's : 0

CDISCPILOT01

USUBJID

NA's : 0

01-701-1015, 01-701-1023, 01-701-1028,

SUBJID

NA's : 0

1,001 1,040 1,081 1,130 1,181 1,220 1,271 1,310 1,361 1,400 1,448

SITEID

NA's : 0

701 702.0 704.0 706.0 708.2 710 711.8 713.0 715.4 717.218

No of observation 254 / 254

254 Rows & 48 Columns

Show 10 entries

| STUDYID      | USUBJID     | SUBJID | SITEID | SITEGR1 | ARM     | TRT01P  | TRT01PN | TRT01A  | TRT01AN | TRTSDT     | TRTEDT     | TRTDUR |
|--------------|-------------|--------|--------|---------|---------|---------|---------|---------|---------|------------|------------|--------|
| CDISCPILOT01 | 01-701-1015 | 1015   | 701    | 701     | Placebo | Placebo | 0       | Placebo | 0       | 2014-01-02 | 2014-07-02 | 182    |
| CDISCPILOT01 | 01-701-1023 | 1023   | 701    | 701     | Pla     |         |         |         |         |            |            |        |
| CDISCPILOT01 | 01-701-1028 | 1028   | 701    | 701     | Xai Hig |         |         |         |         |            |            |        |

1 Upload Data

2 Review Your Data

3 Data Filter

4 Visualize

5 Publish

AVGDD

NA's : 0

0 23 45 67 89 111 133 155 177 199 212

CUMDOSE

NA's : 0

0 1,000 2,200 4,400 6,600 8,800 11,000 12,200 15,417

AGE

NA's : 0

01 55 60 83 87 71 75 79 83 87 89

AGEGR1

NA's : 0

65-80, <65, >80

AGEGR1N

NA's : 0

0 1 2 3 4 5 6 7 8 9

RACEN

NA's : 0

1 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6

AMERICAN INDIAN OR ALASKA NATIVE

BLACK OR AFRICAN AMERICAN

WHITE

AMERICAN INDIAN OR ALASKA NATIVE

## Summary

| STUDYID                                                                                      | USUBJID                                                                                      | SUBJID                                                                                 | SITEID                                                                                       |
|----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| CDISCPILOT01:170                                                                             | Length:170<br>Class :character<br>Mode :character                                            | Min. :1001<br>1st Qu.:1086<br>Median :1190<br>Mean :1204<br>3rd Qu.:1314<br>Max. :1448 | Min. :701.0<br>1st Qu.:704.0<br>Median :708.0<br>Mean :708.3<br>3rd Qu.:713.0<br>Max. :718.0 |
| SITEGR1                                                                                      | ARM                                                                                          | TRT01P                                                                                 | TRT01PN                                                                                      |
| Min. :701.0<br>1st Qu.:704.0<br>Median :709.0<br>Mean :731.5<br>3rd Qu.:716.0<br>Max. :900.0 | Length:170<br>Class :character<br>Mode :character                                            | Length:170<br>Class :character<br>Mode :character                                      | Min. : 0.00<br>1st Qu.: 0.00<br>Median : 0.00<br>Mean :26.68<br>3rd Qu.:54.00<br>Max. :54.00 |
| TRT01A                                                                                       | TRT01AN                                                                                      | TRTSDT                                                                                 | TRTEDT                                                                                       |
| Length:170<br>Class :character<br>Mode :character                                            | Min. : 0.00<br>1st Qu.: 0.00<br>Median : 0.00<br>Mean :26.68<br>3rd Qu.:54.00<br>Max. :54.00 | Length:170<br>Class :character<br>Mode :character                                      | Length:170<br>Class :character<br>Mode :character                                            |
| TRTDUR                                                                                       | AVGDD                                                                                        | CUMDOSE                                                                                | AGE                                                                                          |
| Min. : 2.0<br>1st Qu.: 56.0<br>Median :166.0<br>Mean :124.3<br>3rd Qu.:183.0<br>Max. :212.0  | Min. : 0.00<br>1st Qu.: 0.00<br>Median : 0.00<br>Mean :26.68<br>3rd Qu.:54.00<br>Max. :54.00 | Min. : 0<br>1st Qu.: 0<br>Median : 0<br>Mean :2642<br>3rd Qu.:4481<br>Max. :11448      | Min. :51.00<br>1st Qu.:70.00<br>Median :77.00<br>Mean :75.44<br>3rd Qu.:82.00<br>Max. :89.00 |
| AGEGR1                                                                                       | AGEGR1N                                                                                      | AGEU                                                                                   | RACE                                                                                         |
| Length:170<br>Class :character                                                               | Min. :1.000<br>1st Qu.:2.000                                                                 | Length:170<br>Class :character                                                         | Length:170<br>Class :character                                                               |

# Data visualization:

Graph Designer

Powered by  

1 Upload Data2 Review Your Data3 Data Filter4 Visualize5 Publish

Feedback ⓘ ⓘ ⓘ ⓘ

Layers

One of the key ideas behind `ggplot2` is to build complex graphs — one layer at a time. Thus, if your graph has two layers, the first layer is applied then the second will be on applied on top of it.

If available, use the `+ Add Layer` and `- Remove Layer` buttons below to interactively **add** and **customize** layers and **remove** them, respectively.

Once satisfied, navigate to the fifth step — **Publish**.

AestheticsGeometricScale Fill/Color ManualFacet/Split byAnnotate

TransformationLabels/Theme

Aesthetic(s) Mapping

Aesthetic mappings describe how variables in the data are mapped to visual properties (aesthetics) of **geoms**. Aesthetic mappings can be applied to all layers using this tab or in individual layers using the other tabs such as **geom**.

x

ARM

y

AGE

scale

NONE

+

-

color

ARM

fill

ARM

AGE

ARM

PlaceboXanomeline Low DoseXanomeline High Dose

ARM  
Placebo  
Xanomeline Low Dose  
Xanomeline High Dose

Graph Designer

Powered by  

1 Upload Data2 Review Your Data3 Data Filter4 Visualize5 Publish

Feedback ⓘ ⓘ ⓘ ⓘ

Layers

One of the key ideas behind `ggplot2` is to build complex graphs — one layer at a time. Thus, if your graph has two layers, the first layer is applied then the second will be on applied on top of it.

If available, use the `+ Add Layer` and `- Remove Layer` buttons below to interactively **add** and **customize** layers and **remove** them, respectively.

Once satisfied, navigate to the fifth step — **Publish**.

AestheticsGeometricScale Fill/Color ManualFacet/Split byAnnotate

TransformationLabels/Theme

Geom Layer(s) Mapping

For a given geom layer, you can start annotating your graph by selecting a geom type and configure by selecting attribute(s) which will dynamically populated widgets.

Depending on the attribute, you can set the attribute by mapping a column from the input data or specify a **value**.

Layer 1

geom\_boxplot

Attribute(s) Selector

Layer 2

geom\_jitter

Attribute(s) Selector

+ Add Layer

- Remove Layer

AGE

ARM

PlaceboXanomeline Low DoseXanomeline High Dose

ARM  
Placebo  
Xanomeline Low Dose  
Xanomeline High Dose

Show Plot



# Visualization:

1 Upload Data

2 Review Your Data

3 Data Filter

4 Visualize

5 Publish

Feedback

Aesthetics

Geometric

Scale Fill/Color Manual

Facet/Split by

Annotate

Transformation

Labels/Theme

### Facet/Split By

With faceting, you can make multi-panel plots and control how the scales of one panel relate to the scales of another. This is useful if you have a single variable with many levels and want to arrange the plots in a more space efficient manner.

Faceting generates small multiples each showing a different subset of the data, based on a separation category variable with multiple unique levels. Small multiples are a powerful tool for exploratory data analysis: you can rapidly compare patterns in different parts of the data and see whether they are the same or different.

You can control how the ribbon is wrapped into a grid with `ncol`, `nrow`, `as.table` and `dir`. `ncol` and `nrow` control how many columns and rows (you only need to set one). `as.table` controls whether the facets are laid out like a table (TRUE), with highest values at the bottom-right, or a plot (FALSE), with the highest values at the top-right. `dir` controls the direction of wrap: horizontal or vertical.

Read the official documentation for `facet_wrap` by clicking [here](#)

Attribute(s) Selector

facets

facets

SEX

Confirm

1 Upload Data

2 Review Your Data

3 Data Filter

4 Visualize

5 Publish

Feedback

- Use the **caption** to provide information about the data source
- Tags** can be used for adding identification tags to differentiate between multiple plots

Plot Title

Age Distribution by Cohort

X-axis Label

ARM

Y-axis Label

AGE

Caption

Enter text

### Theme Customization

Themes are a powerful way to customize the non-data components of your plots: i.e. titles, labels, fonts, background, grid lines, and legends. Themes can be used to give plots a consistent customized look.

Overall Plot Theme Selection

Plot Theme

Classic

Plot Title

Title Color

#000000

Title Size

20

Horizontal Adjust

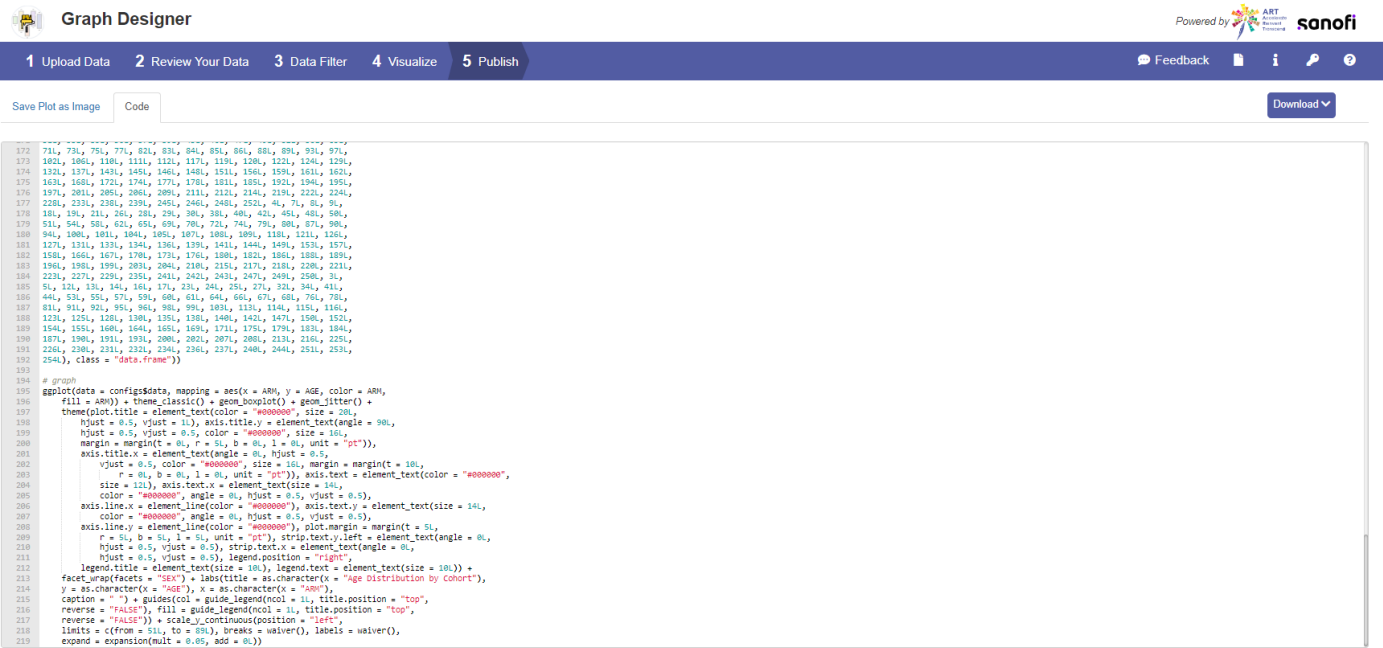
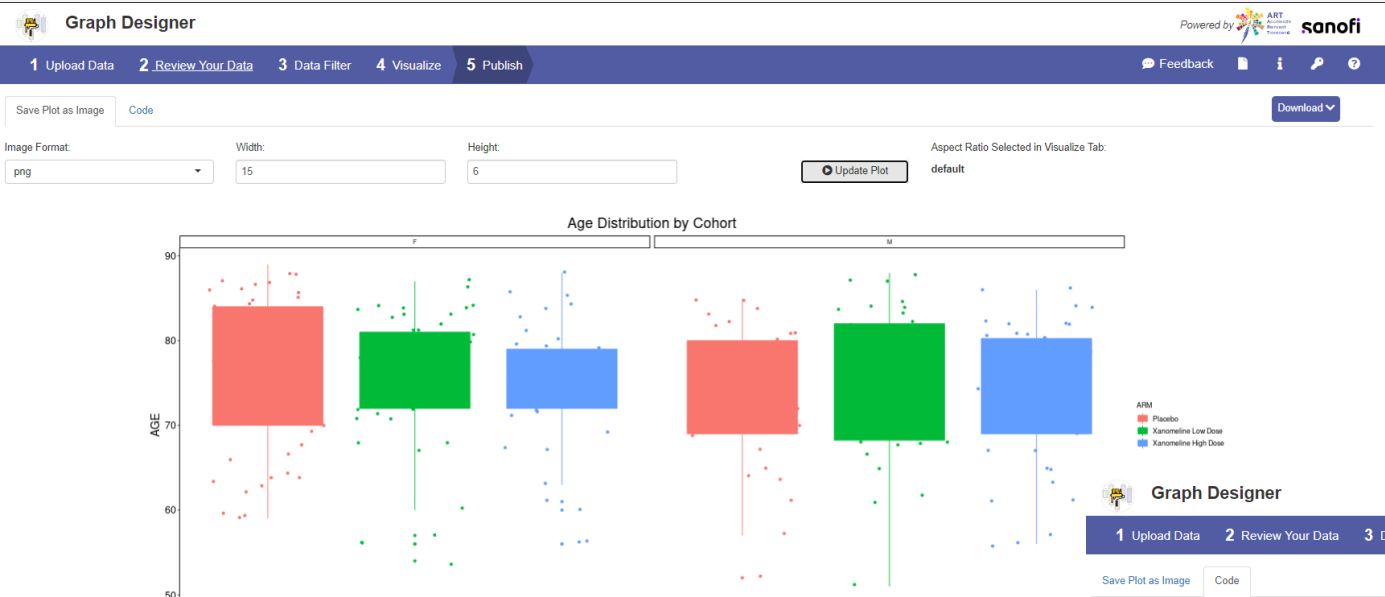
0.5

Vertical Adjust

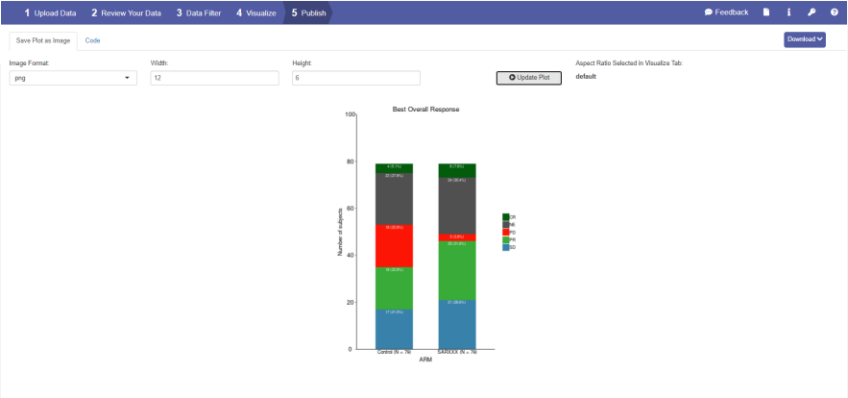
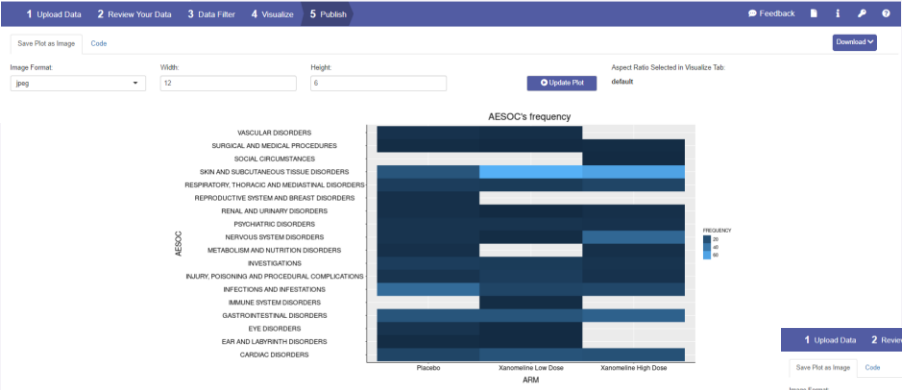
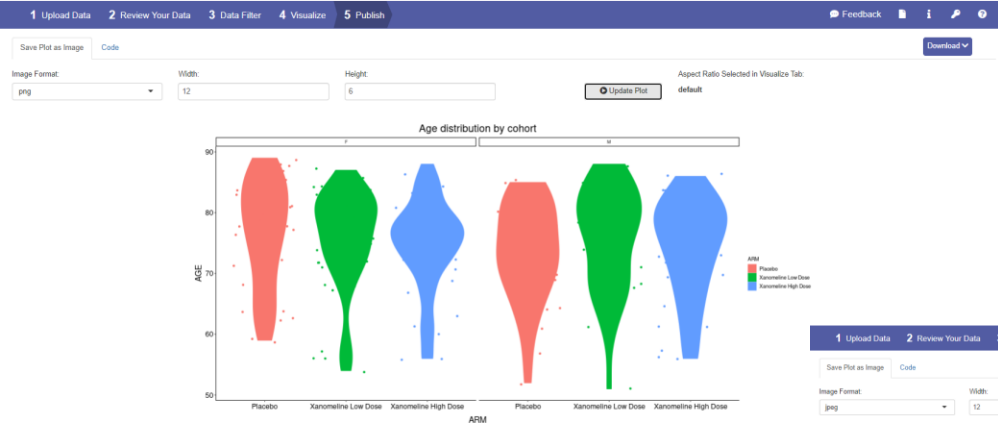
1

Update Plot

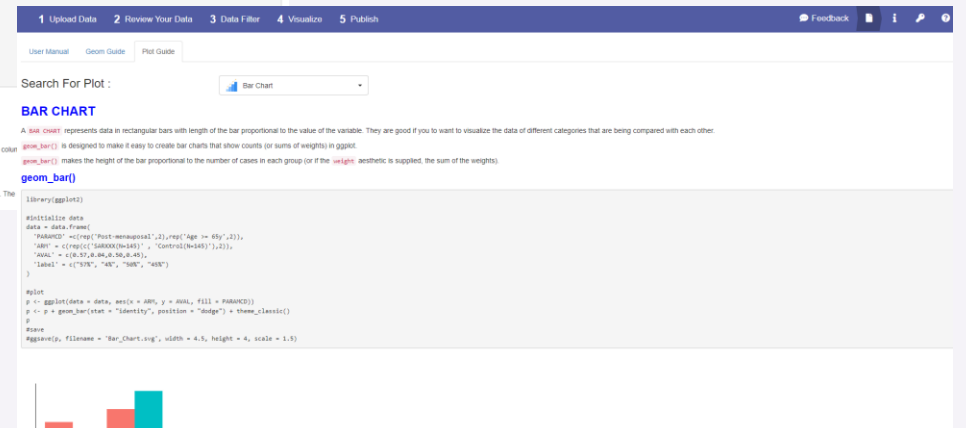
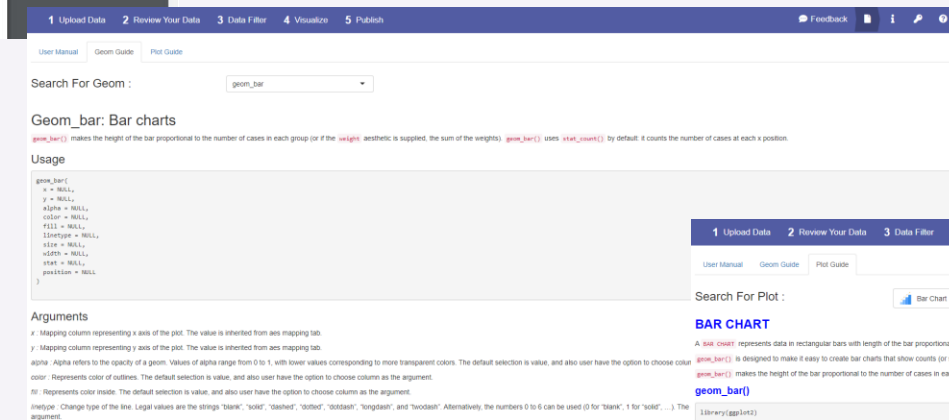
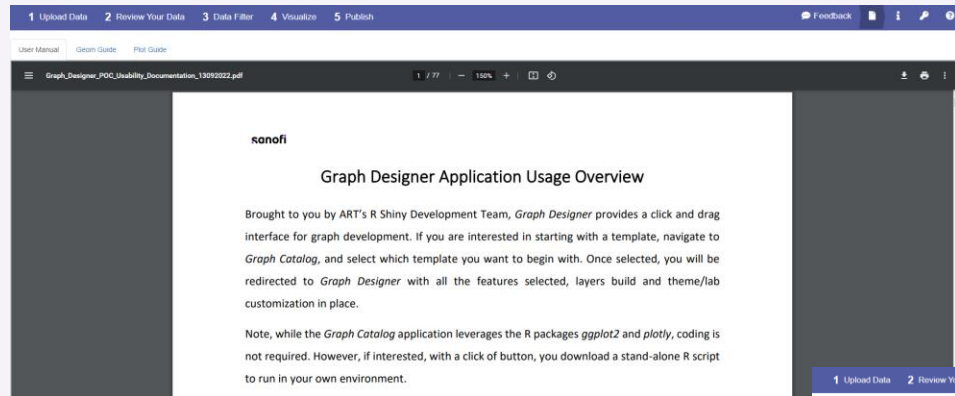
# Finalizing the plot:



# Different Plots:



## User guide and plot information:



# Salient features:

- The Graph Designer provides a user-friendly interface for graph generation, where users can click and select to create publication quality graphs from the scratch without prior knowledge of any programming language.
- This application is a large wrapper around the R package **ggplot2**, which is a system for declaratively creating graphs, based on "**The Grammar of Graphics**" and behind the scenes, R code is compiled and executed.
- Graph Designer is a self-sufficient tool (Software as a service (**SaaS**)) to generate publication quality figures. It allows users to turn data into coherent, visually immersive, and interactive insights without spending time on programming or on other trivial applications.
- Graph Designer is also regenerative. Respective code is generated along with visualization, which covers all the process from data input, data filtration to visualization. Graphs and respective code with the input data can be downloaded, code can be reused, adapted/validated to ensure reproducibility, and maximize generalizability.

# Acknowledgements



**Xiaodong**



**Benjamin**



**Andre Couturier**



**Freeman**

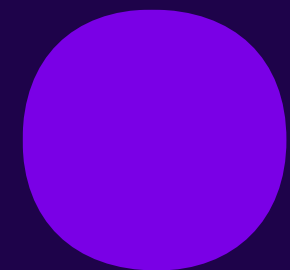


**Mayank\***



**Sujeet\***

**Demo**





sonofi



Thank You



**sanofi**