



INTEGO
GROUP

Paper DV02

CATCH THEM WHEN YOU CAN... FIX THEM

Anna Panasenko



AGENDA

- Rationale
- Tools
- Examples
 - 1) Scatter plot
 - 2) Swimmer plot
 - 3) Line plot
- Summary



RATIONALE

- Easier to perceive
- Time-saving
- More (correct) data = robust results and complete picture
- Less workarounds
- Ideas for checks





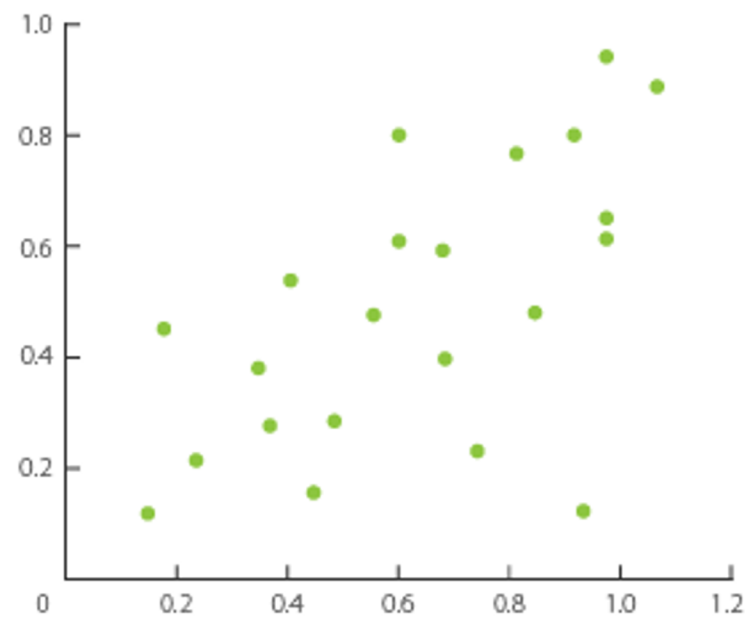
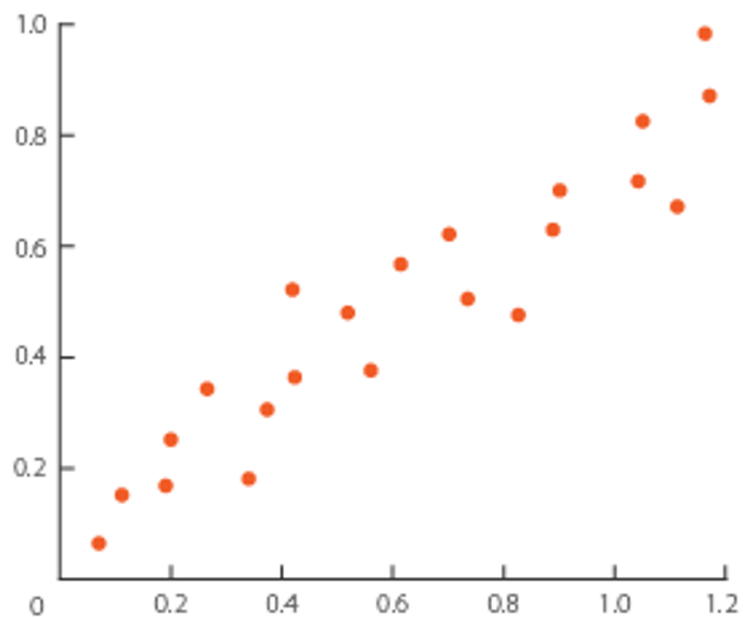
TOOLS

- SAS
 - + everyone knows
 - + built-in tools (ODS Graphics Designer, snippets)
 - may be complicated
- R
 - + easy to customise
 - + no preprocessing
 - + options (base plot functions vs ggplot2 package)



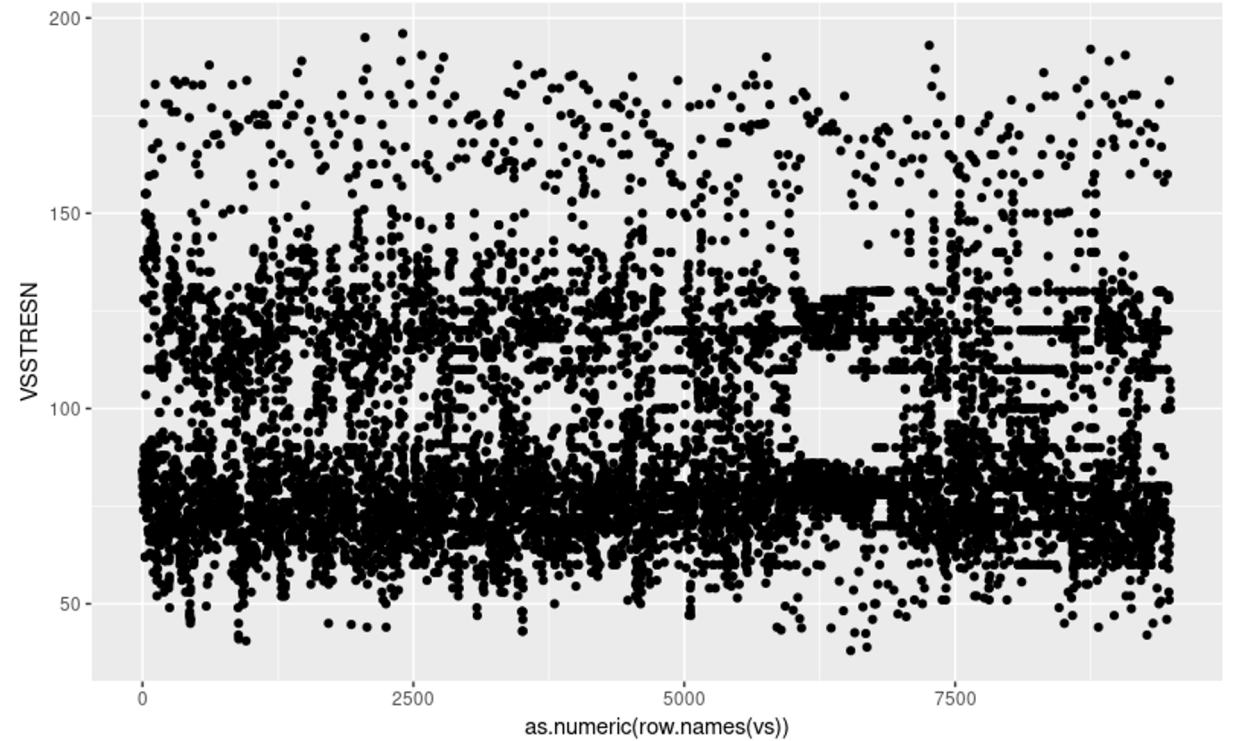
EXAMPLE I

SCATTER PLOT



SCATTER PLOT

- Single variable characterisation
- Numeric test results
- Shows outliers



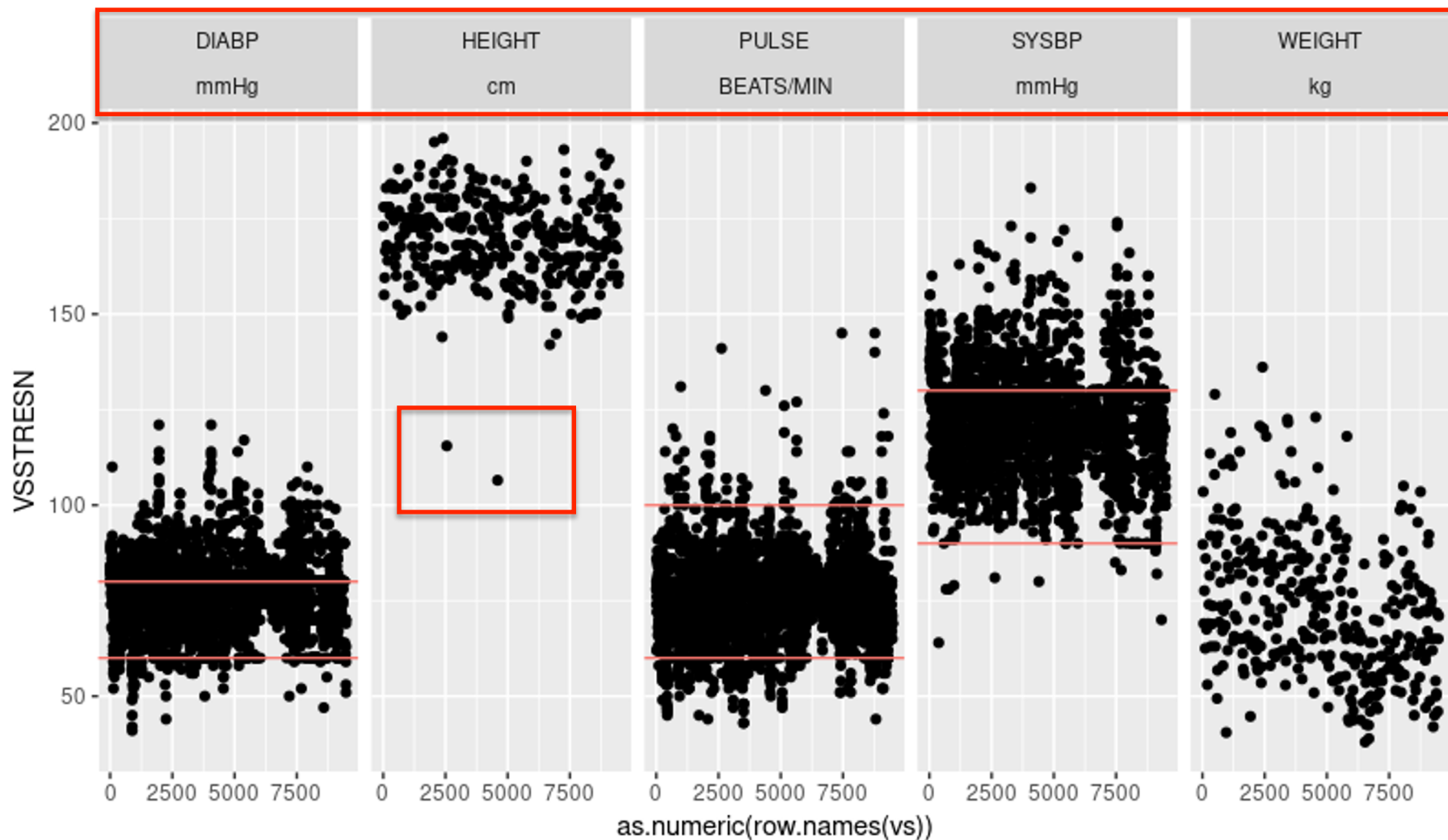
```
# request vital signs scatter plot  
with index x-axis and standard result on y-axis  
scatter <- ggplot(data = vs,  
                  aes(x = as.numeric(row.names(vs)), y = VSSTRESN)) +  
  geom_point()
```

SCATTER PLOT

- Split by test and unit and add reference ranges

```
# request vital signs scatter plot
with index x-axis and standard result on y-axis
scatter +
#split by test and unit
facet_grid(cols = vars(VSTESTCD, VSSTRESU)) +
# lower reference limit
geom_hline(data = vs, aes(yintercept = VSSTRNLO, color = 'red'), show.legend =
FALSE) +
# upper reference limit
geom_hline(data = vs, aes(yintercept = VSSTRNHI, color = 'red'), show.legend =
FALSE)
```

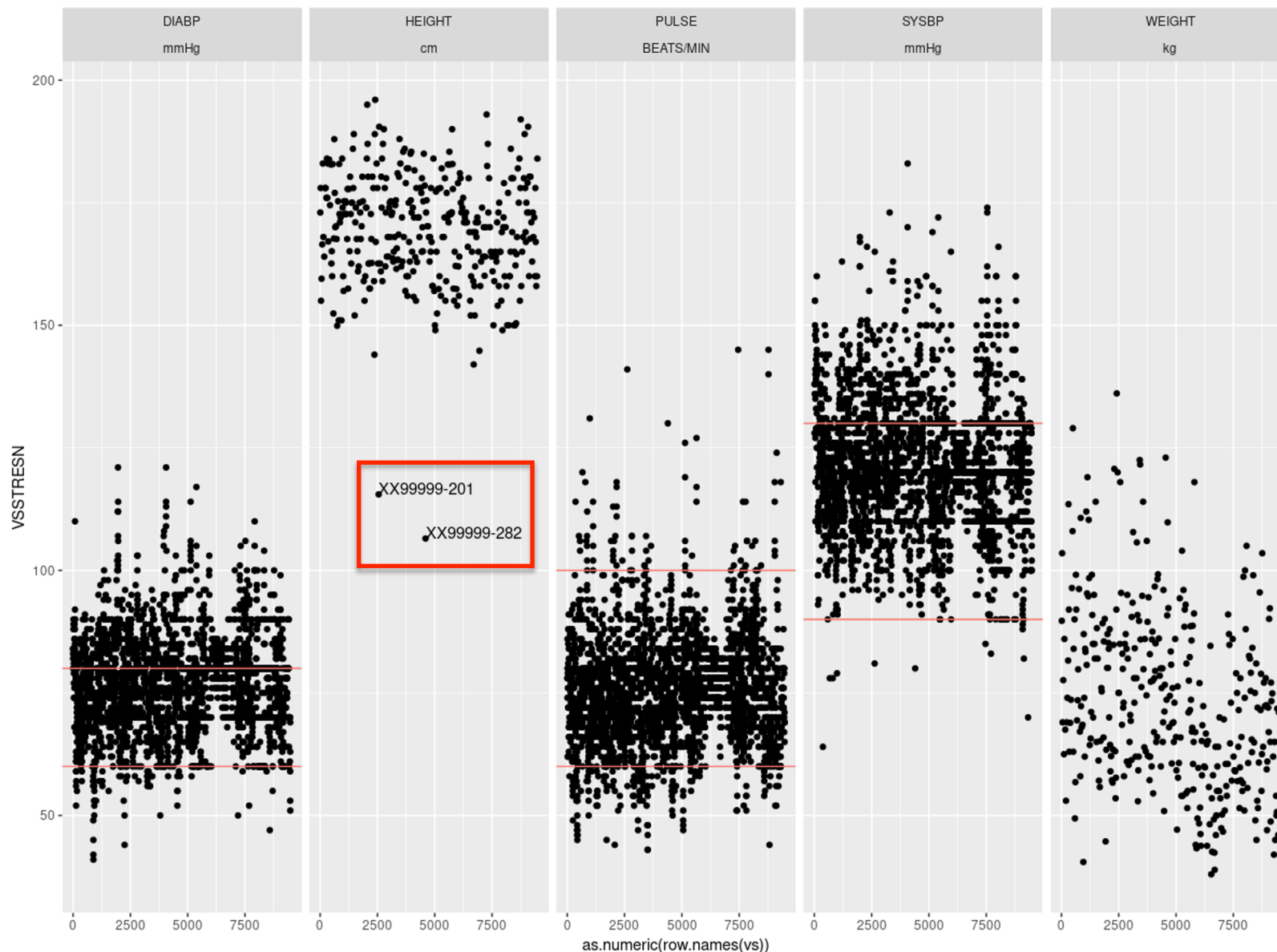
SCATTER PLOT



SCATTER PLOT

- Add subject ID

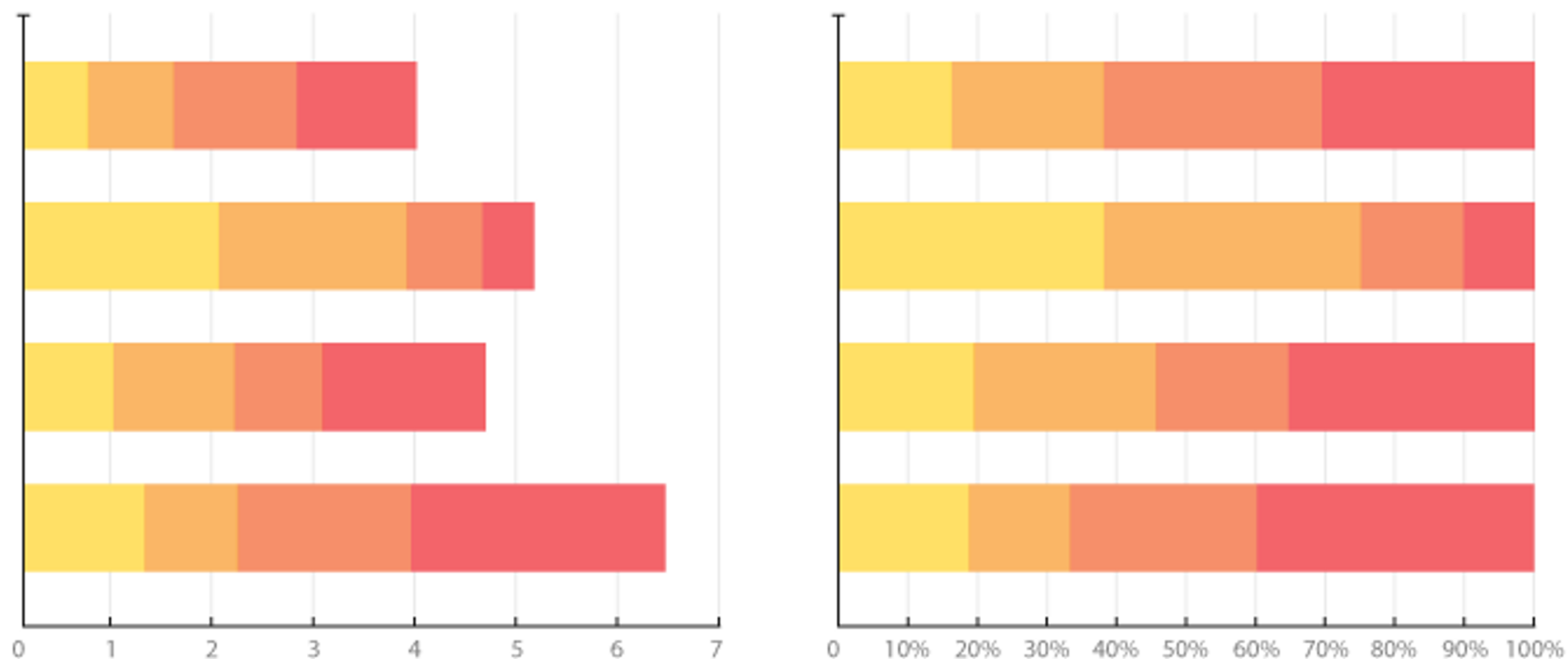
```
# previous graph
scatter +
# add USUBJID to observations
of interest
geom_text(aes(label = ifelse(
  (VSSTRESN < 125 &
   VSTESTCD == 'HEIGHT')
  USUBJID,
  '')),
  hjust = 0,
  vjust = 0)
```





EXAMPLE II

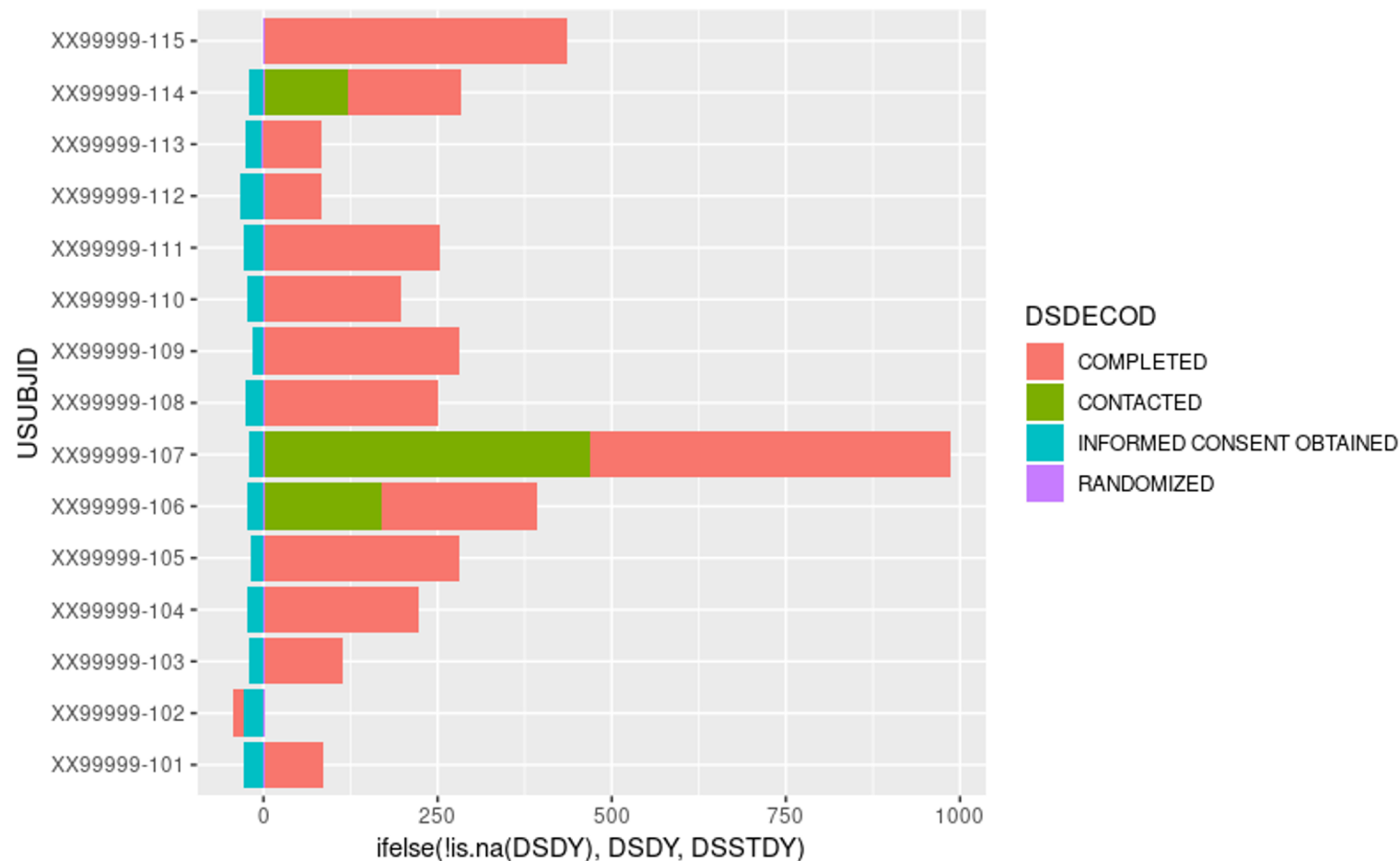
SWIMMER PLOT



SWIMMER PLOT

- Sequential order
- Disposition data

```
# request disposition plot
ggplot(ds,
       aes(USUBJID,
           ifelse(!is.na(DSDY), DSDY, DSSTDY))) +
# add bar plot layer
geom_bar(stat = 'identity', position = 'identity',
        aes(fill = DSDECOD)) +
#rotate
coord_flip()
```



SWIMMER PLOT

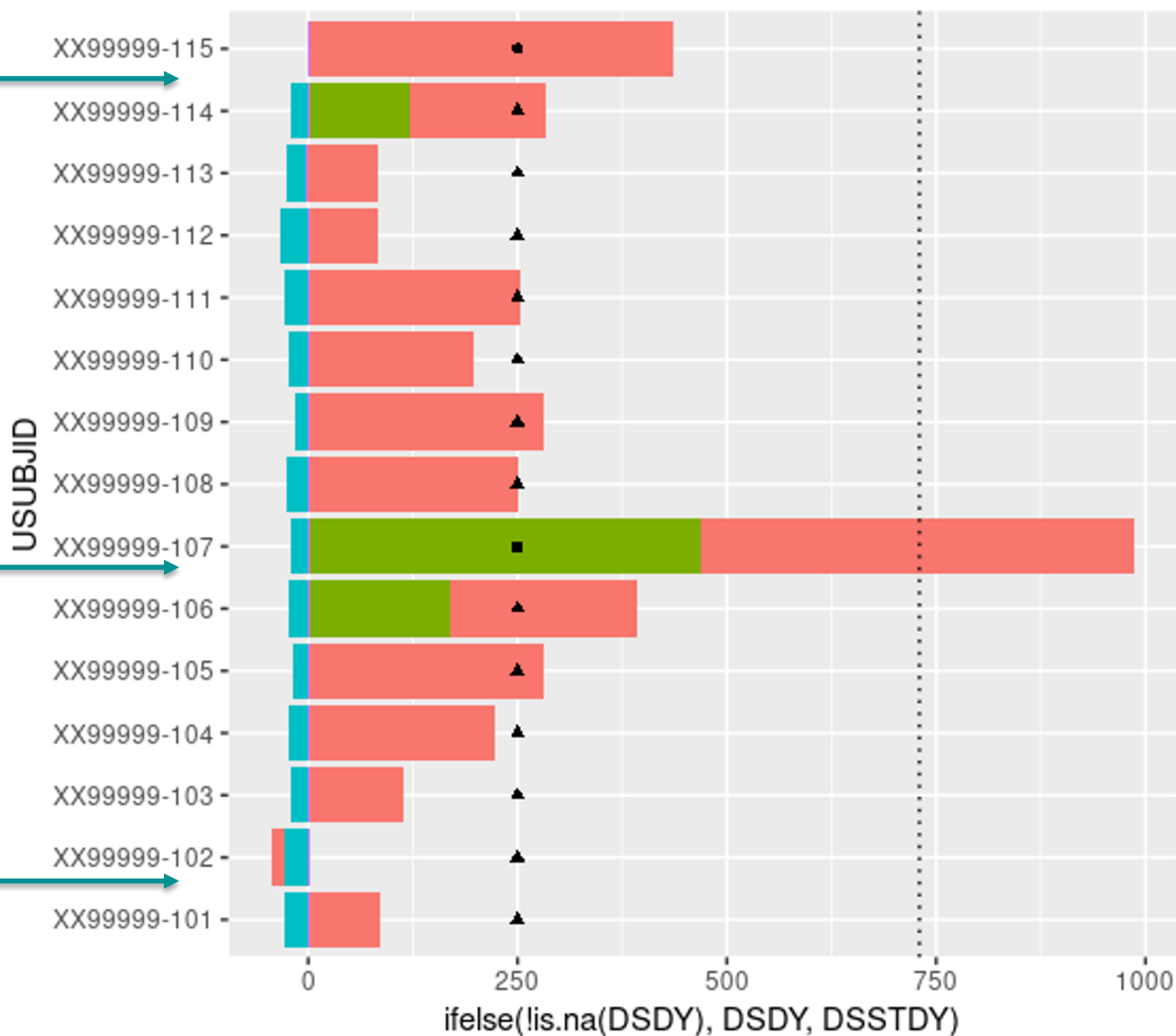
```
# request disposition plot
ggplot(ds,
       aes(USUBJID,
           ifelse(!is.na(DSDY), DSDY, DSSTDY))) +
# add bar plot layer
geom_bar(stat = 'identity', position = 'identity', aes(fill = DSDECOD)) +
# rotate
coord_flip() +
# reference line with max duration of the study
geom_hline(aes(yintercept = 730),
           linetype = 'dotted',
           show.legend = FALSE)
# study drug completion status
geom_point(subset(ds, DSSCAT == 'STUDY DRUG COMPLETION/EARLY DISCONTINUATION'),
           aes(x = USUBJID, y = 250, shape = DSDECOD),
           position = position_dodge(width = 1))
```

SWIMMER PLOT

No informed consent date

Duration is too long

Completion before IC



DSDECOD

- COMPLETED
- LACK OF EFFICACY
- WITHDRAWAL BY SUBJECT

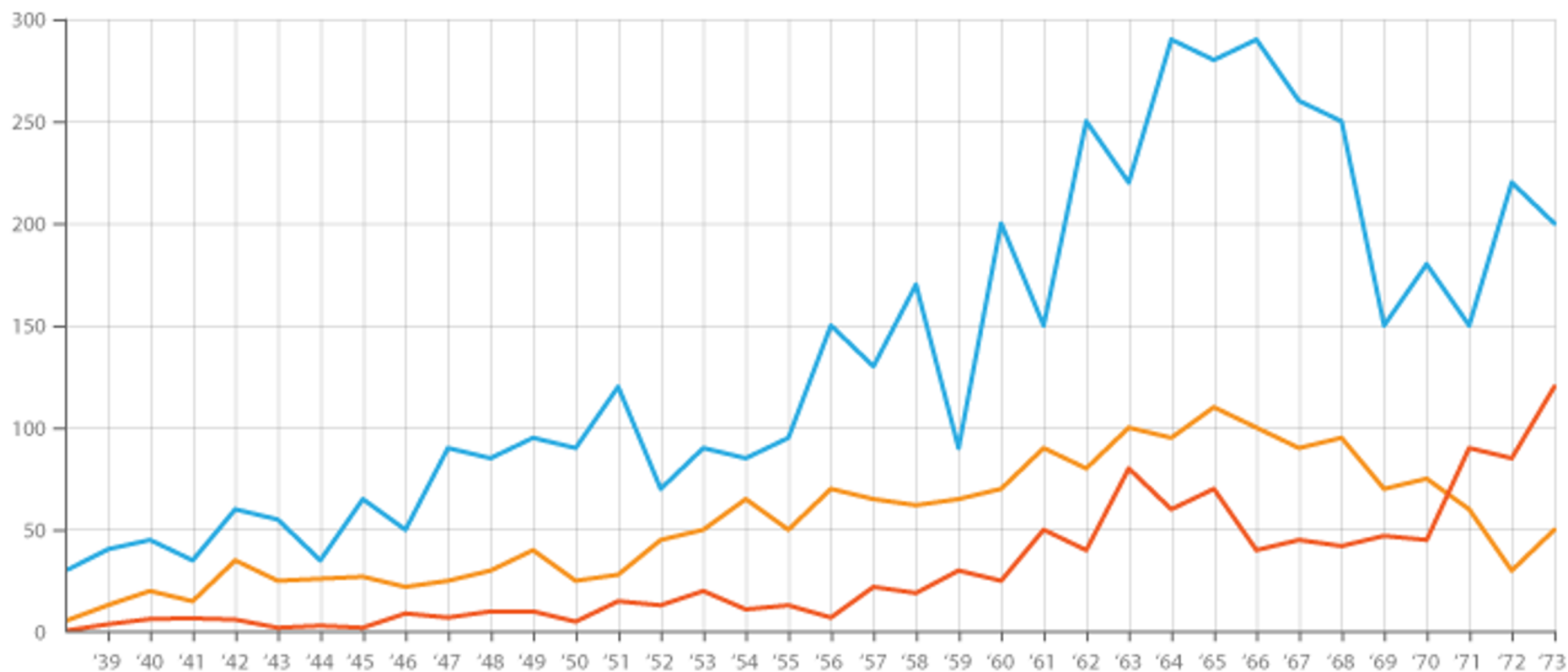
DSDECOD

- COMPLETED
- CONTACTED
- INFORMED CONSENT OBTAINED
- RANDOMIZED



EXAMPLE III

LINE PLOT

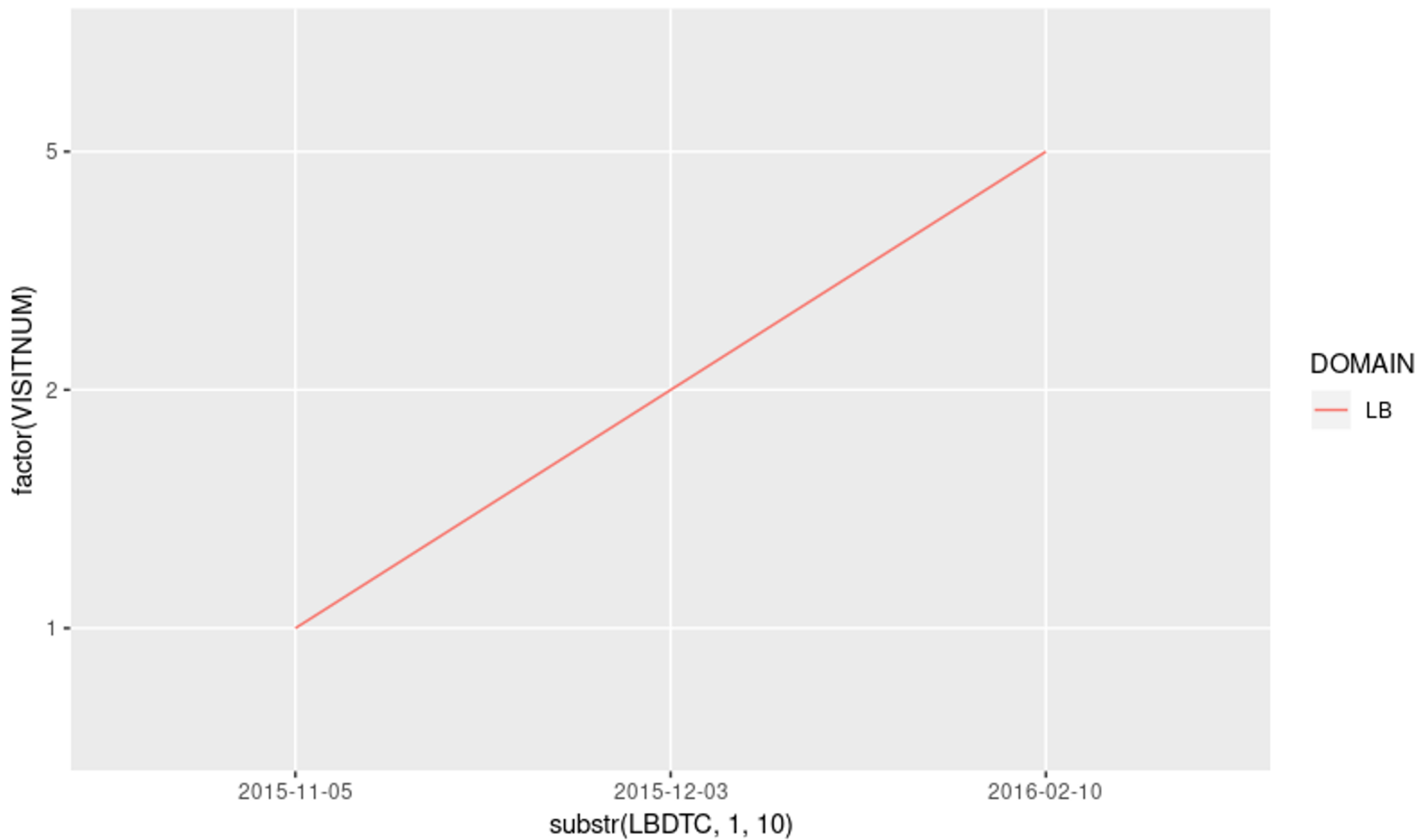


LINE PLOT

- Time series
- Individual patient plot of dates across domains

```
# request empty plot
ggplot() +
# add a line for LB domain
geom_line(data = lb,
          aes(substr(LBDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1)
```

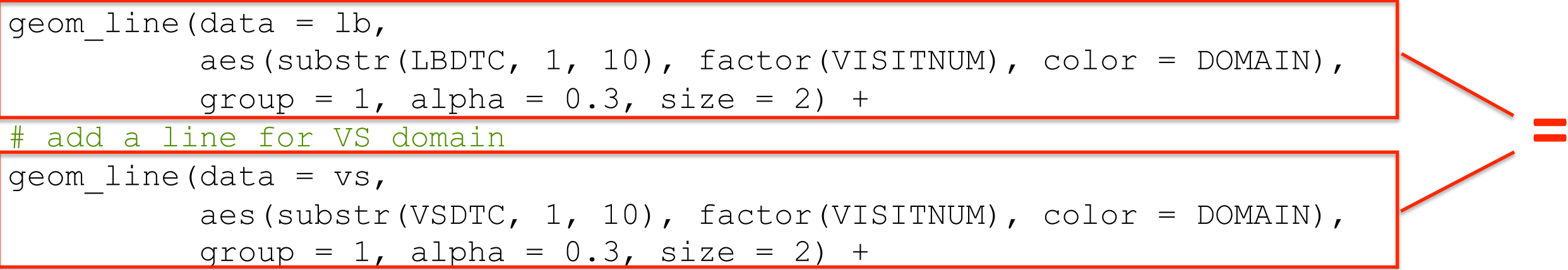
LINE PLOT



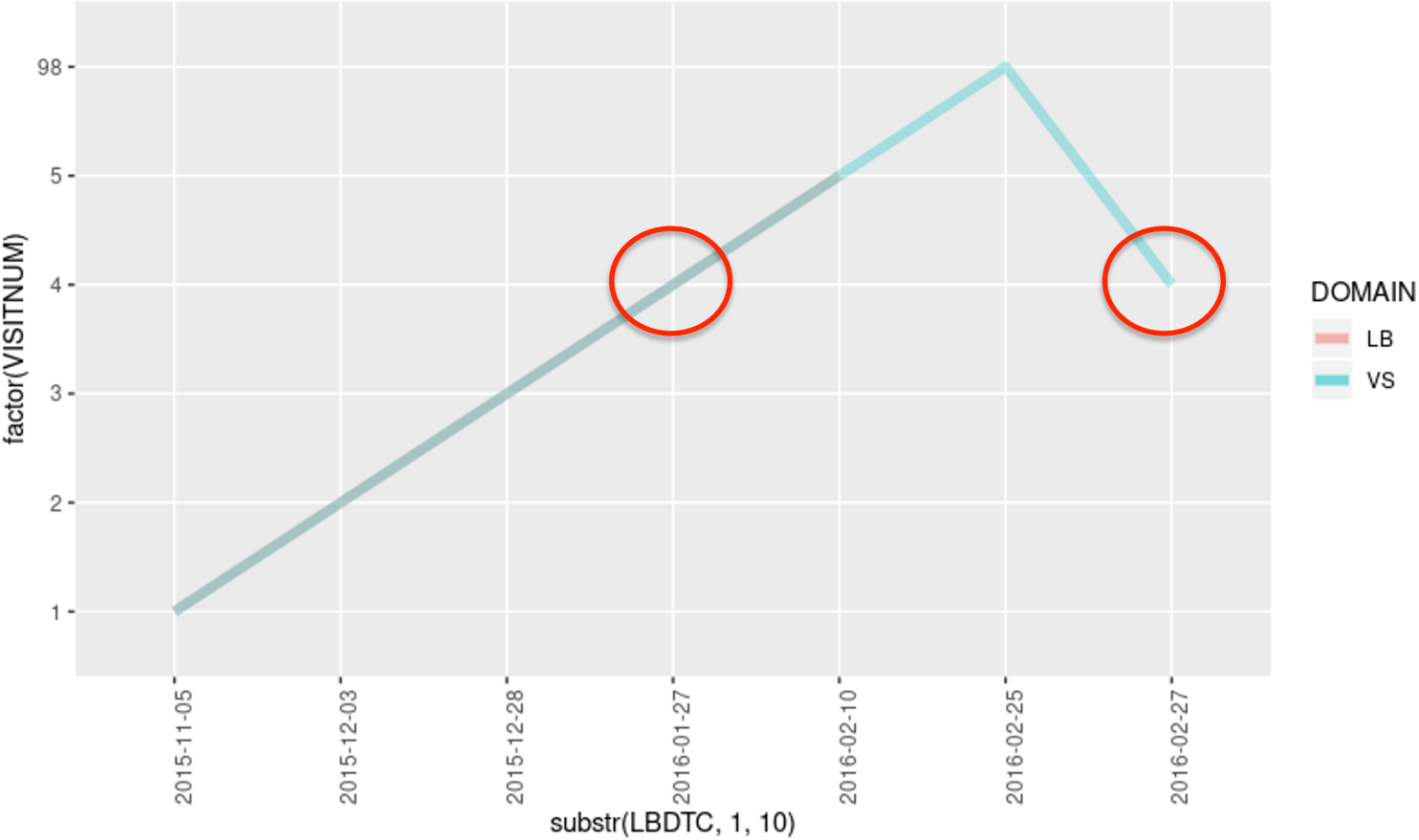
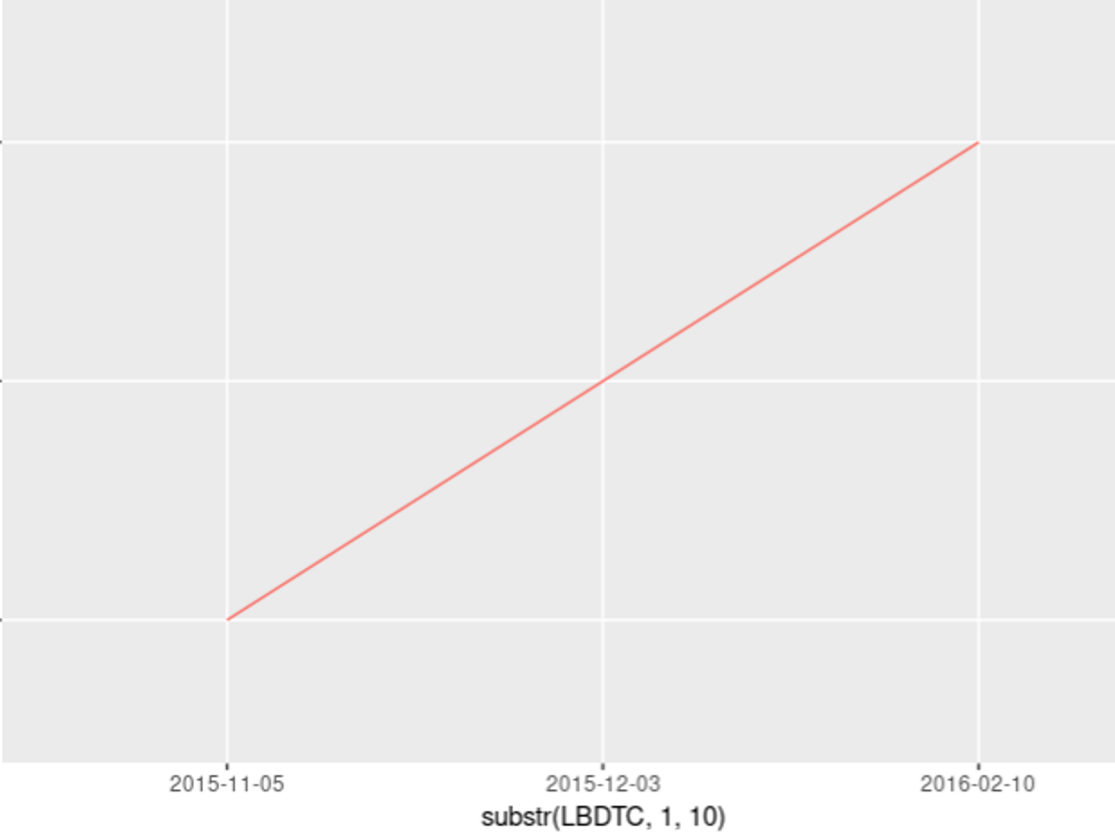
LINE PLOT

- Add one more dataset
- Rotate dates on the x-axis

```
# request empty plot
ggplot() +
# add a line for LB domain
geom_line(data = lb,
          aes(substr(LBDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1, alpha = 0.3, size = 2) +
# add a line for VS domain
geom_line(data = vs,
          aes(substr(VSDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1, alpha = 0.3, size = 2) +
# rotate x-axis tick labels
theme(axis.text.x = element_text(angle = 90))
```

A diagram illustrating the concatenation of two code blocks. Two red rectangular boxes are shown, one above the other. The top box contains the first two lines of the R code (comment and geom_line for LB domain). The bottom box contains the next two lines (comment and geom_line for VS domain). Red lines from the right side of each box converge towards a red equals sign (=) on the right, indicating that the code in the two boxes is to be combined.

LINE PLOT

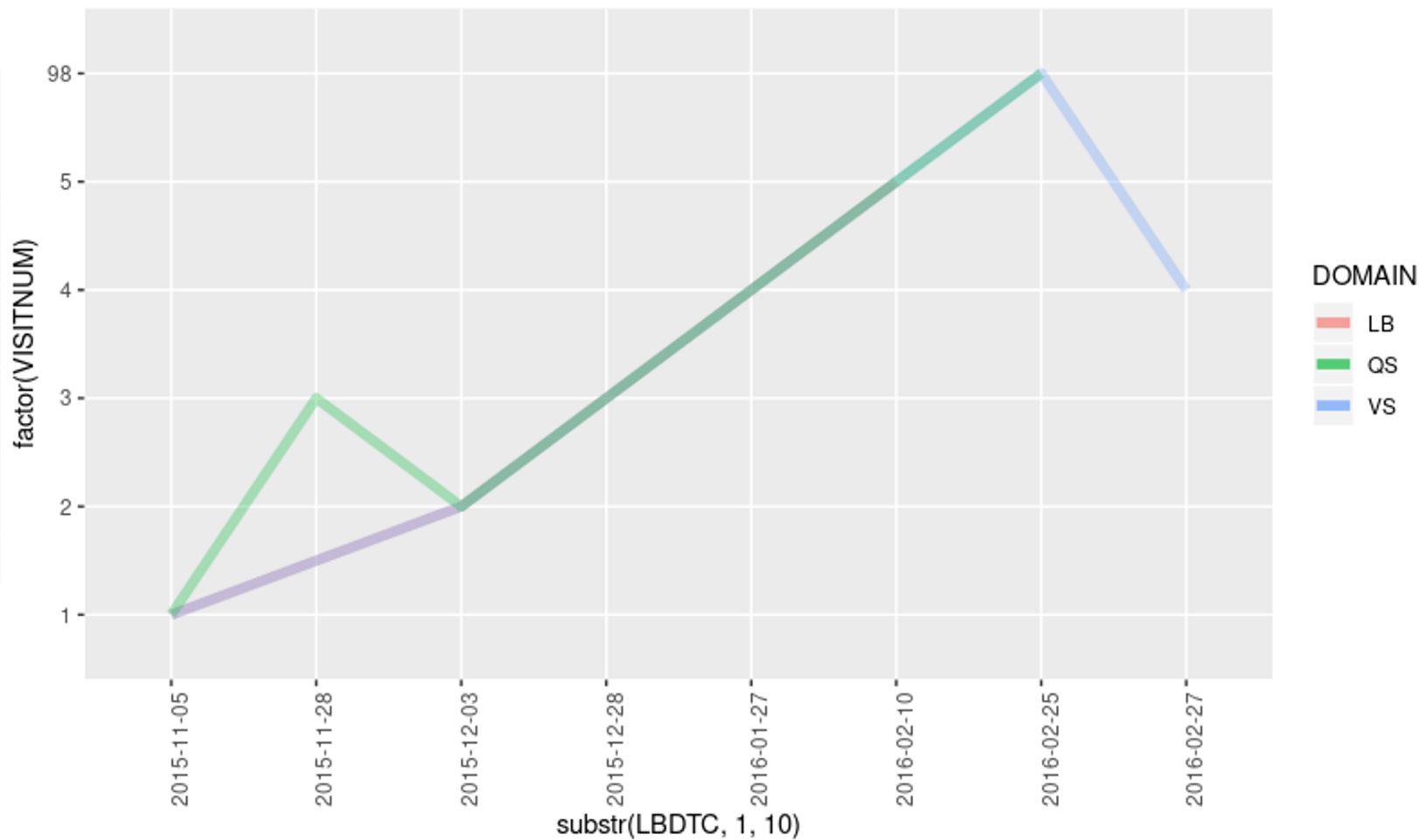
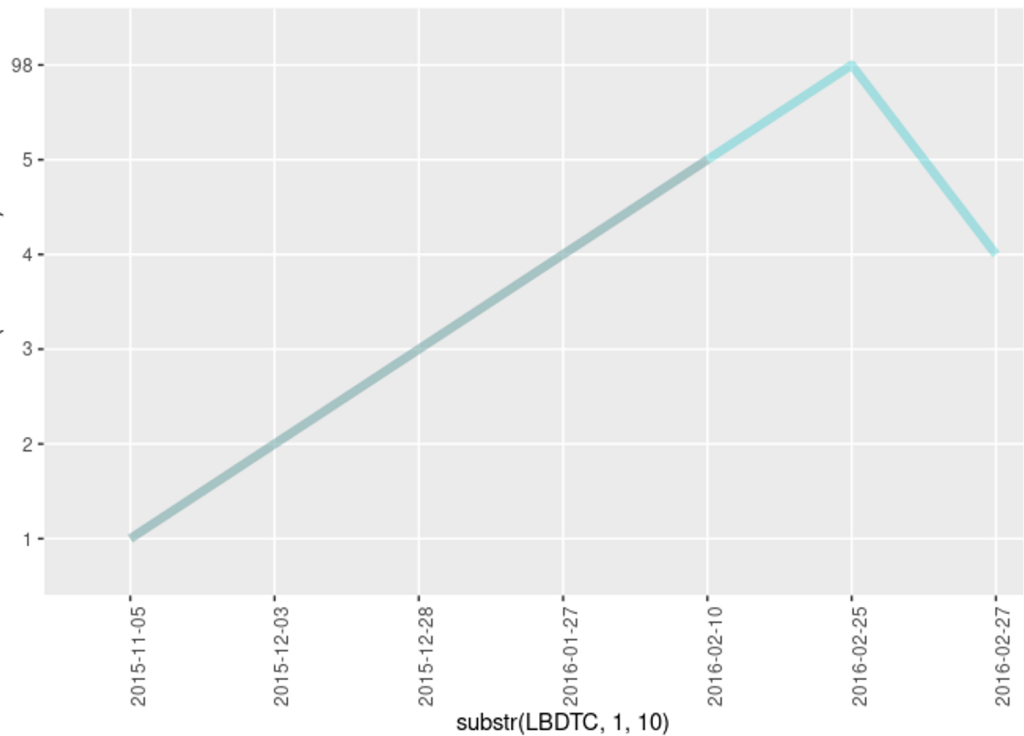


LINE PLOT

```
#request empty plot
ggplot() +
#add a line for LB domain
geom_line(data = lb,
          aes(substr(LBDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1, alpha = 0.3, size = 2) +
#add a line for VS domain
geom_line(data = vs,
          aes(substr(VSDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1, alpha = 0.3, size = 2) +
#add a line for QS domain
geom_line(data = qs,
          aes(substr(QSDTC, 1, 10), factor(VISITNUM), color = DOMAIN),
          group = 1, alpha = 0.3, size = 2) +
#rotate x-axis tick labels
theme(axis.text.x = element_text(angle = 90))
```

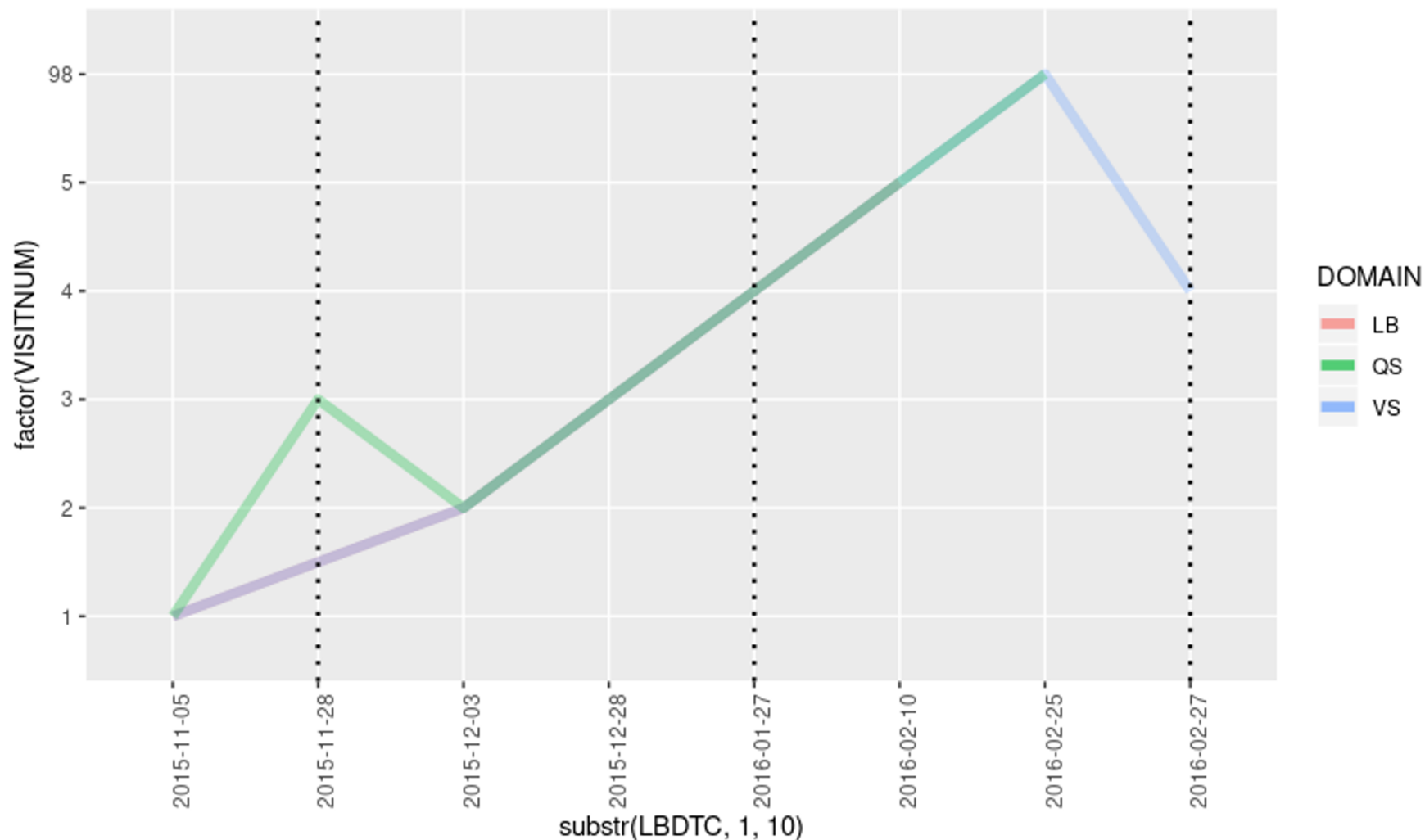


LINE PLOT



LINE PLOT

- Add lines and shapes from other domains, e.g. randomisation, AEs etc.



SUMMARY

- Graphs make it easier to explore your data
- No need for fancy graphs
- Proactively catching data issues rather than at TLG stage
- ‘You can’t always get what you want’



THANK YOU

Anna Panasenko
Intego Group LLC
19 Hromadyanska Street
Kharkiv 61057, Ukraine
Work Phone: +1 407.512.1006 (ext. 2445)
Email: anna.panasenko@intego-group.com
Web: www.intego-group.com

