



Improvement on Sankey Diagram SAS Macro and Implementation

Alvin Mathew, Teresa Curto & Weiming Du

PHUSE SM02

Agenda



Introduction

Motivation

Existing Solution

Improvement

Next Steps

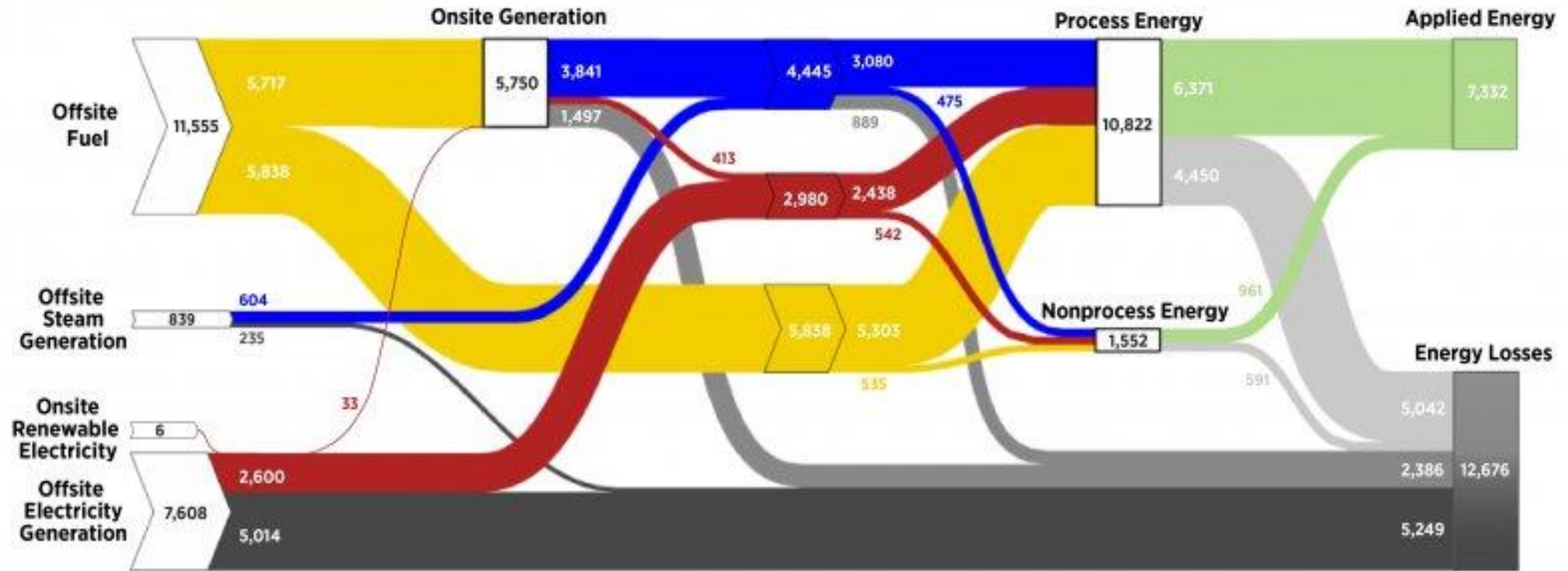
Agenda



Introduction

What is a Sankey diagram?

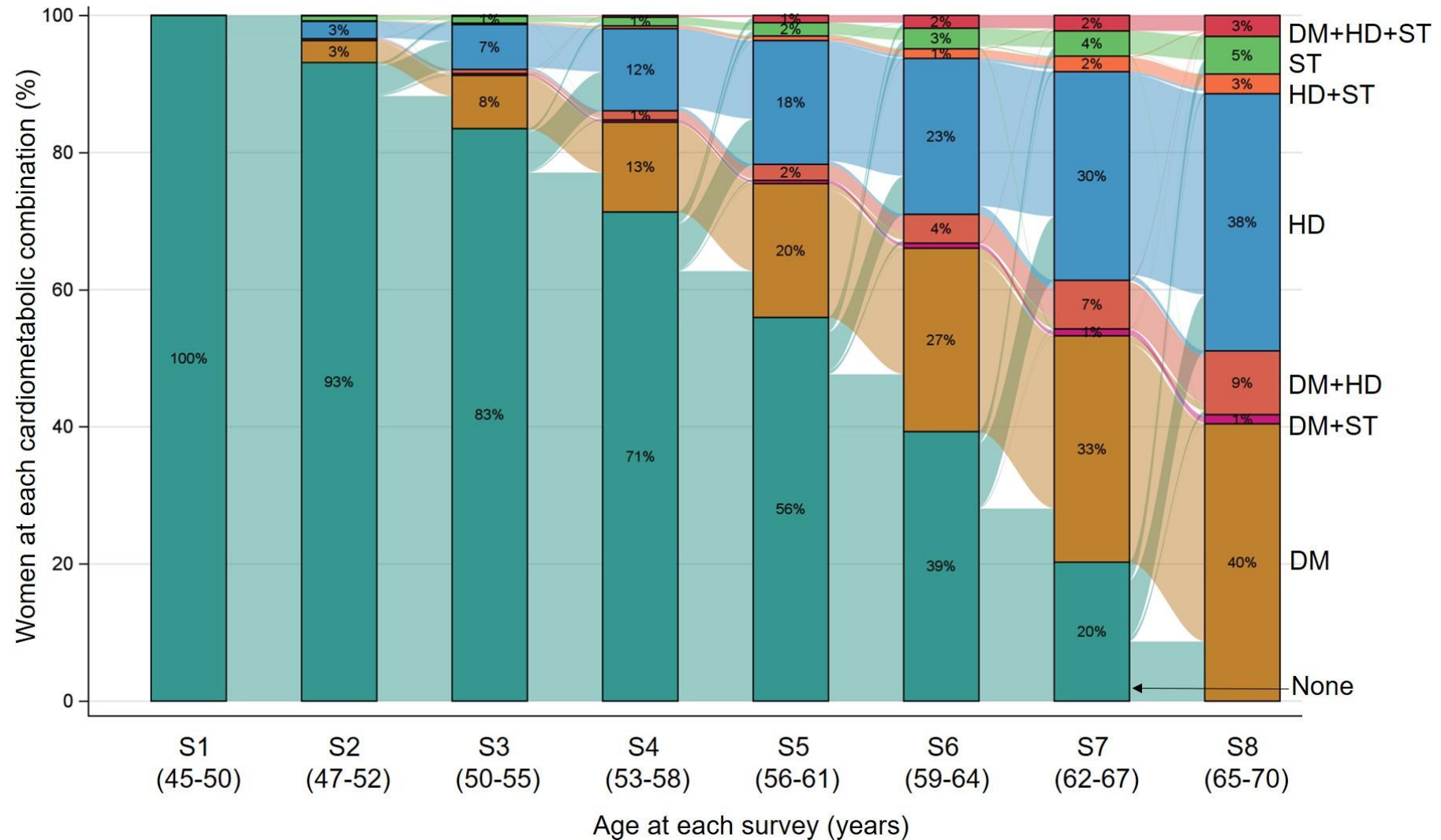
U.S. Manufacturing Sector (TBtu), 2014



LEGEND: Fuel Steam Electricity Applied Energy Offsite Generation and Transmission Losses
Onsite Generation and Distribution Losses End Use Losses

Introduction

Increasing use in healthcare research

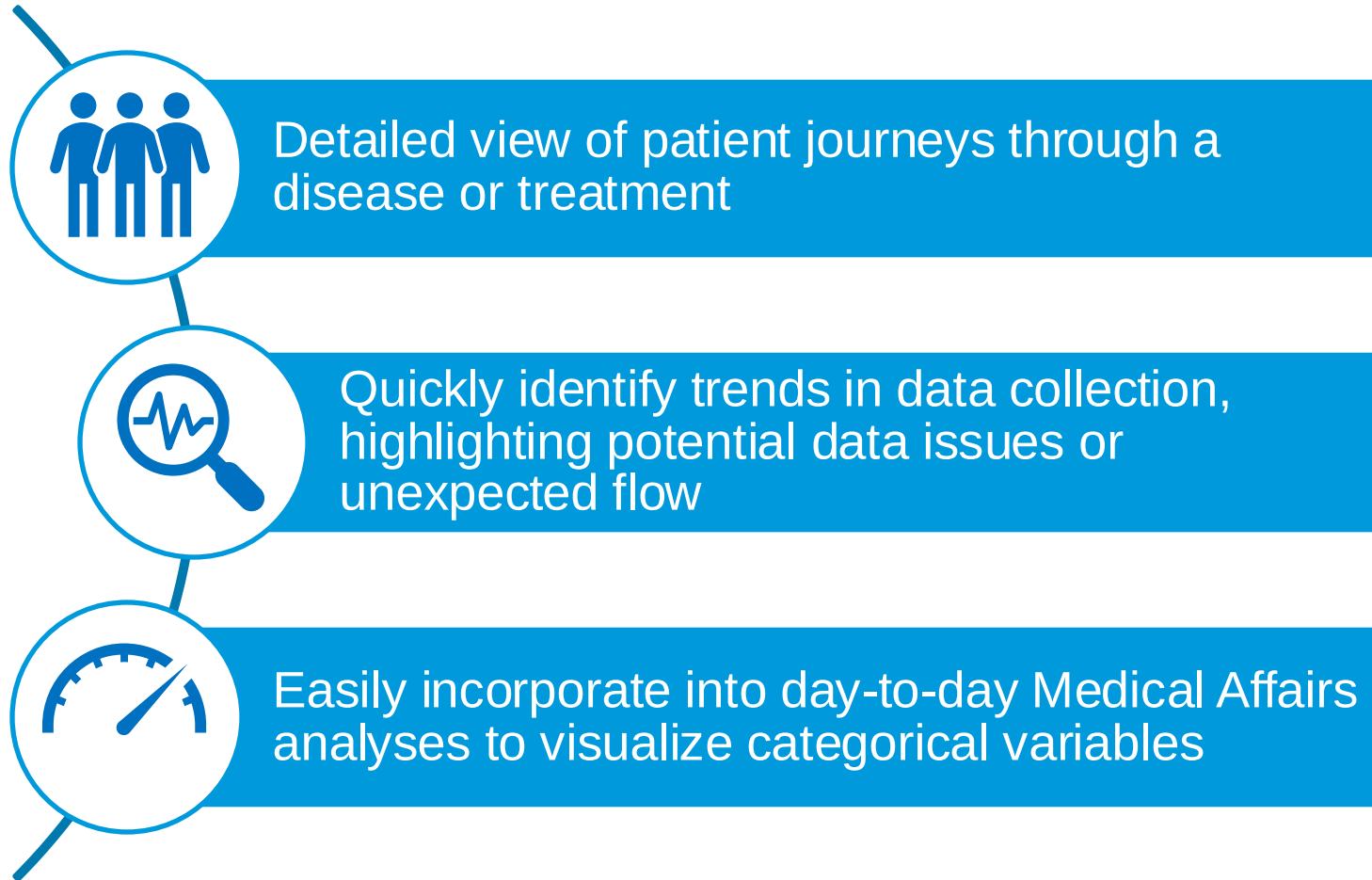


Agenda



Motivation to use

Many benefits with Sankey plots



Agenda

Introduction

Motivation

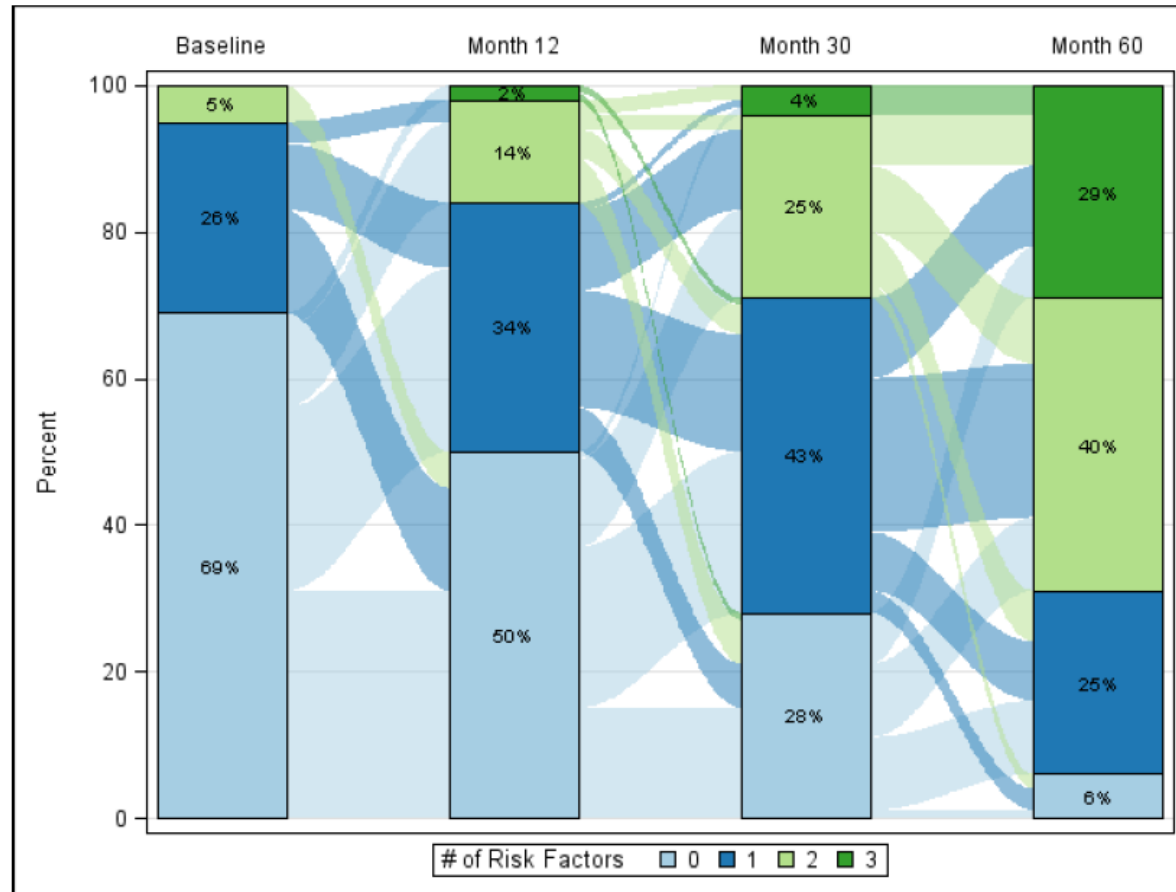
Existing Solution

Improvement

Next Steps

Existing Solution

One approach in SAS



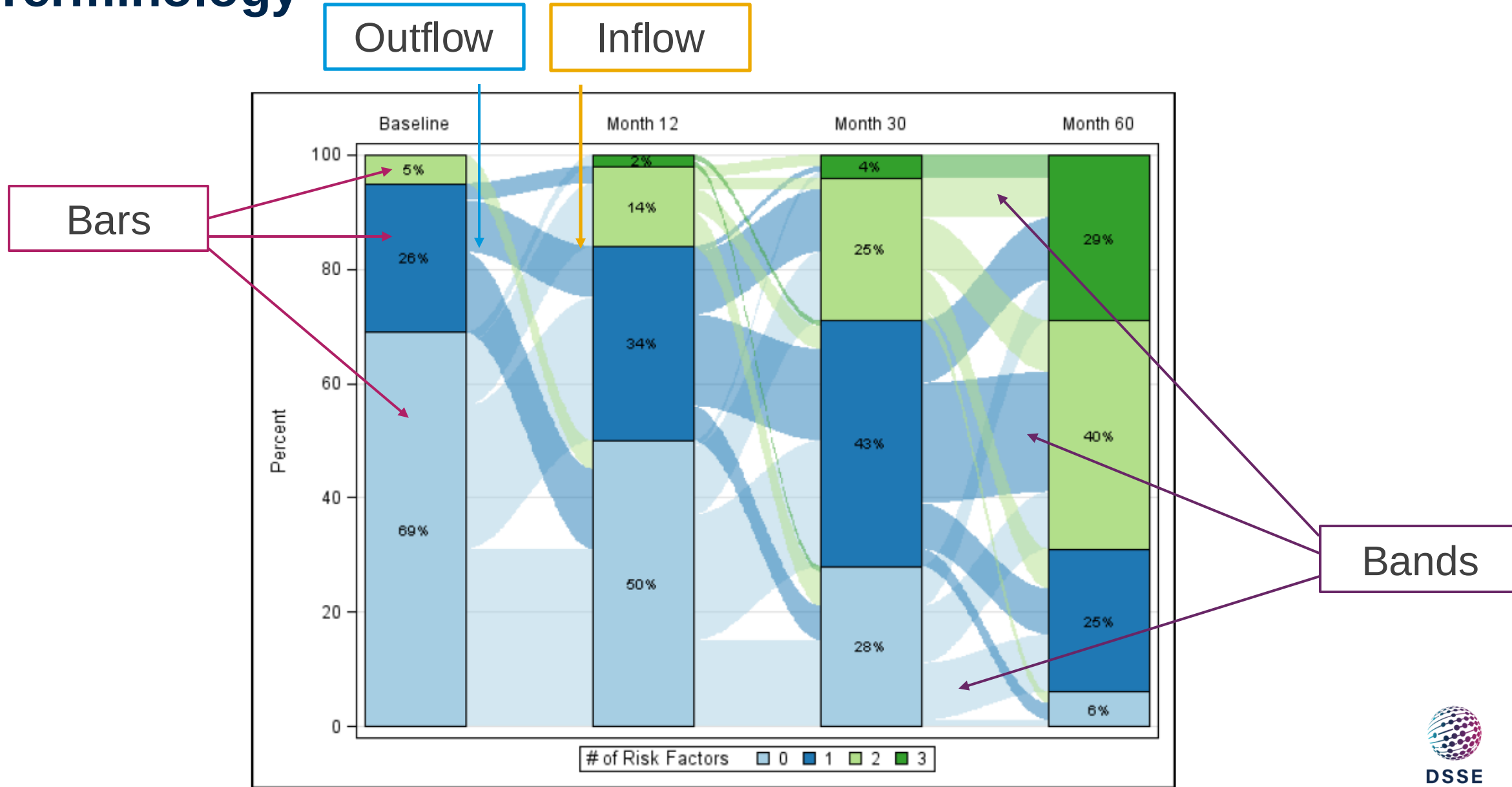
Pro

- Intuitive visualization of the change between categories by visits

Con

- Lose detail of number and % of out-flow to and in-flow from the categories across visits

Terminology



Agenda



Improvements

1

Display outgoing and incoming number or percentages

2

Generate summary table that pairs with Sankey plot

3

Color customization

Display outgoing and incoming number or percentages

Determine location of band annotation

```
if first.link or last.link then do;
scatterband_y&jc=mean(yblow&jc,ybhigh&jc);
if first.link then
  scatterband_x&jc=xt&jc+&band_lbl_pos;
else if last.link then scatterband_x&jc=xt&jc -
  &band_lbl_pos;
tempval='value_to_fill';
scatterband= "scatter x=scatterband_x&jc
y=scatterband_y&jc / x2axis
markerchar=tempval;";
```

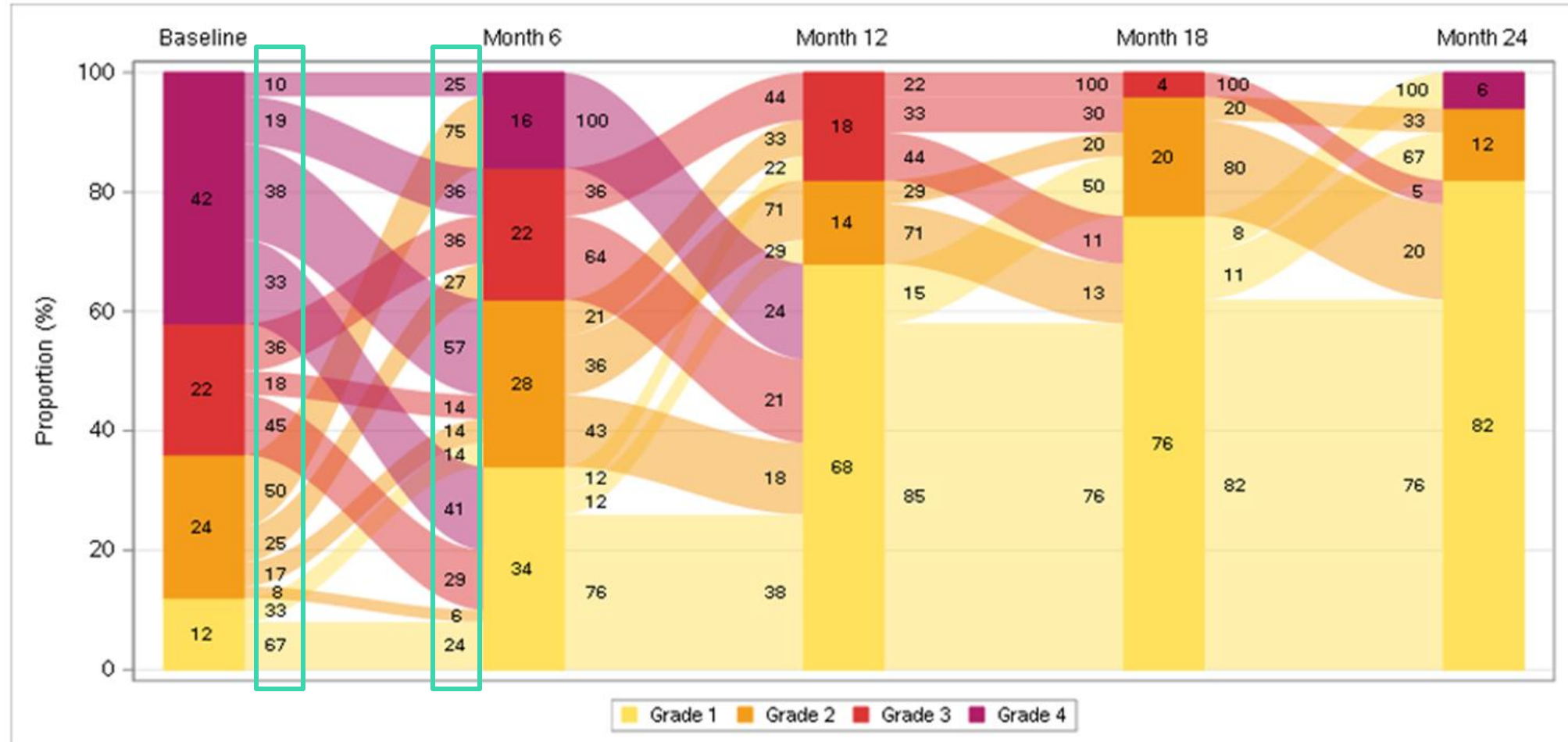
Use band height to find number of patients

```
array low{*} low1-low&n_yvar.;
array high{*} high1-high&n_yvar.;
array den{*} denom1-denom&n_yvar.;
do i = 1 to dim(low);
if nmiss(low{i},ybhigh,high{i})=0 then do;
if . < round(low{i},.0001)+0.01 <=
  round(ybhigh*100,.0001) <=
  round(high{i},.0001)+0.01 then do;
act_ybar=i;
denom=den{i};
leave;
```

Use _highlow dataset to find the cutoffs for the next bar and calculate proportion

```
if ^missing(act_ybar) then do;
perc=put(round(sumpts/denom*100),3.);
ct=strip(put(sumpts,best.);
%if &stat=percent %then
  %do;scatterband2=tranwrd(scatterband,
  'tempval','perc');%end;
%else %if &stat=freq %then
  %do;scatterband2=tranwrd(scatterband,
  'tempval','ct');%end;
end;
```

Display outgoing and incoming number or percentages



Generate summary table that pairs with Sankey plot

Utilize existing information to generate table data

```
/*Calculate table values from macro data*/
outputN=input(resolve('&xy_'||strip(put(x1,best.))||"_"||strip(put(y1,best.))),best.);
inputN=input(resolve('&xy_'||strip(put(x2,best.))||"_"||strip(put(y2,best.))),best.);
rowpct='('||put(round(thickness/outputN*100,.1),5.1)||')';
colpct='('||put(round(thickness/inputN*100,.1),5.1)||')';
count=put(thickness,3.);
tabpct=count||' '||rowpct||' '||colpct;

/*Proc report is built by iterating and storing column data in macro variables*/
colcall&i="('"||strip(_&i.)||" (*ESC*)\brdrb\brdrs\brdrw7' ct&i row&i col&i)";

/*Color information is also stored and called if table color option is triggered*/
rowcolor="if rowtxt='"||strip(value)||"' then call
define(_row_,'style','style={borderbottomcolor='"||strip(fillcolor)||"
borderrightcolor='"||strip(fillcolor)||" background='"||strip(fillcolor)||"'})";
```

Example Table Output

Optional to match table rows to categorical colors

Table 1
Shift Table by Visit

Month 6

Previous Category	Category														
	Grade 1			Grade 2			Grade 3			Grade 4			Total		
	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %
Grade 1	4	(66.7)	(23.5)	2	(33.3)	(14.3)	0	0	0	0	0	0	6	(100.0)	(12.0)
Grade 2	1	(8.3)	(5.9)	2	(16.7)	(14.3)	3	(25.0)	(27.3)	6	(50.0)	(75.0)	12	(100.0)	(24.0)
Grade 3	5	(45.5)	(29.4)	2	(18.2)	(14.3)	4	(36.4)	(36.4)	0	0	0	11	(100.0)	(22.0)
Grade 4	7	(33.3)	(41.2)	8	(38.1)	(57.1)	4	(19.0)	(36.4)	2	(9.5)	(25.0)	21	(100.0)	(42.0)
Total	17	(34.0)	(100.0)	14	(28.0)	(100.0)	11	(22.0)	(100.0)	8	(16.0)	(100.0)	50	(100.0)	(100.0)

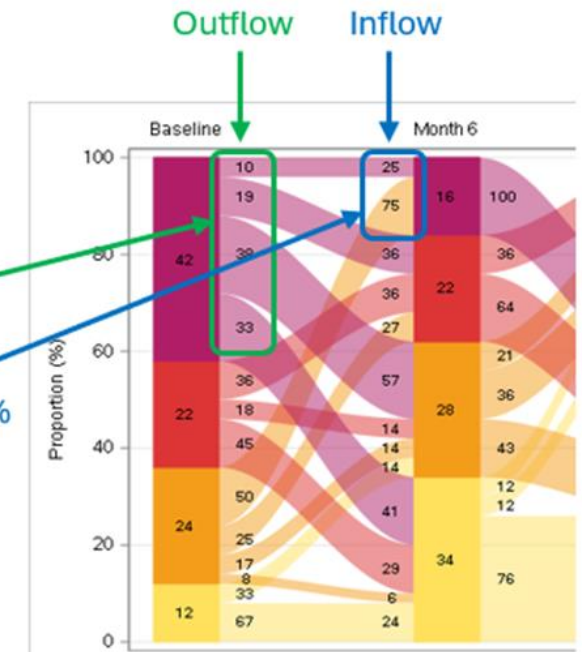
Table and figure relation

Table 1
Shift Table by Visit
Month 6

Previous Category	Grade 1			Grade 2			Grade 3			Grade 4			Total		
	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %	n	Row %	Col %
Grade 1	4	(66.7)	(23.5)	2	(33.3)	(14.3)	0	0	0	0	0	0	6	(100.0)	(12.0)
Grade 2	1	(8.3)	(5.9)	2	(16.7)	(14.3)	3	(25.0)	(27.3)	6	(50.0)	(75.0)	12	(100.0)	(24.0)
Grade 3	5	(45.5)	(29.4)	2	(18.2)	(14.3)	1	(36.4)	(36.4)	0	0	0	11	(100.0)	(22.0)
Grade 4	7	(33.3)	(41.2)	8	(38.1)	(57.1)	4	(19.0)	(36.4)	2	(9.5)	(25.0)	21	(100.0)	(42.0)
Total	17	(34.0)	(100.0)	14	(28.0)	(100.0)	11	(22.0)	(100.0)	8	(16.0)	(100.0)	50	(100.0)	(100.0)

Row %

Col %

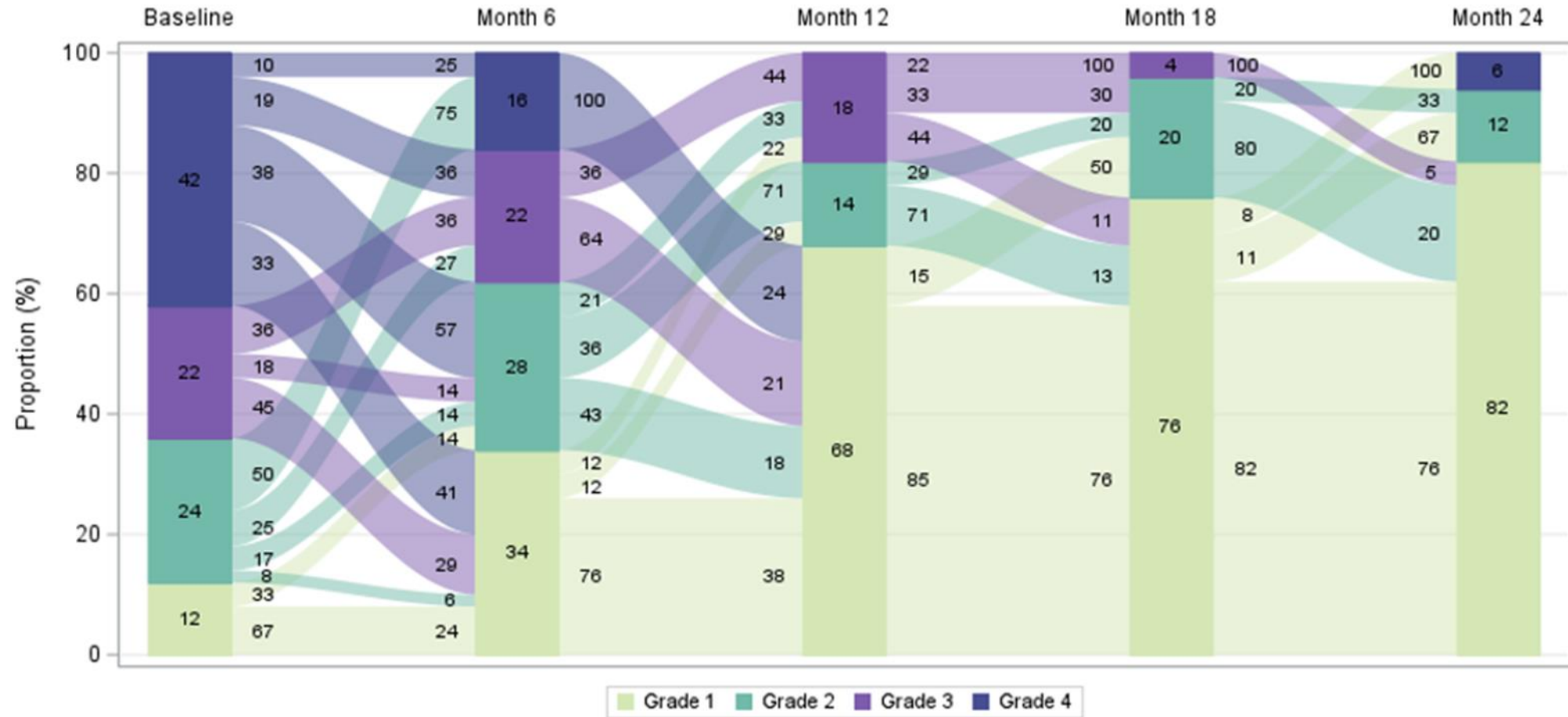


DSSE
CONNECTING THE DOTS

Alnylam

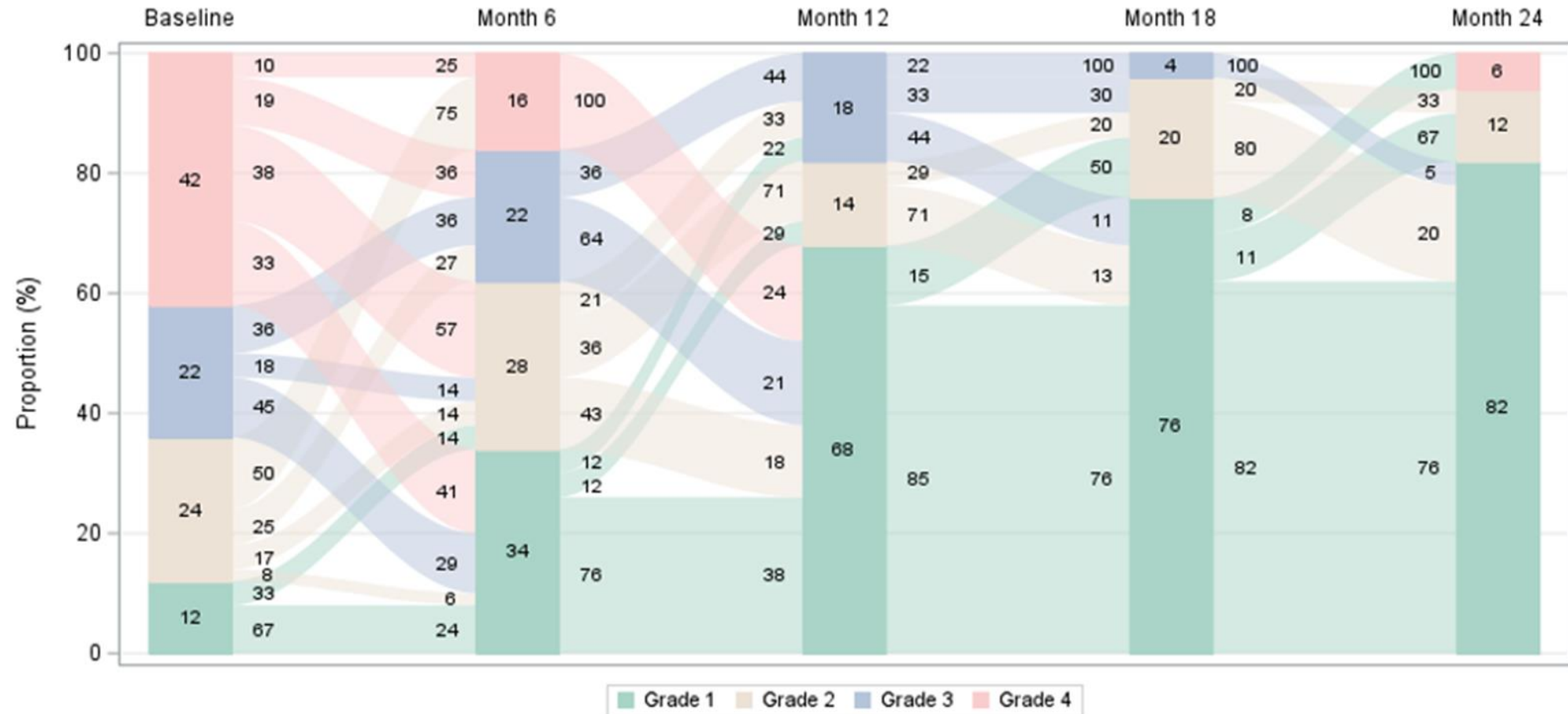
Color Customization

Modified code to accept attribute maps



Color Customization

Change in ATTRID



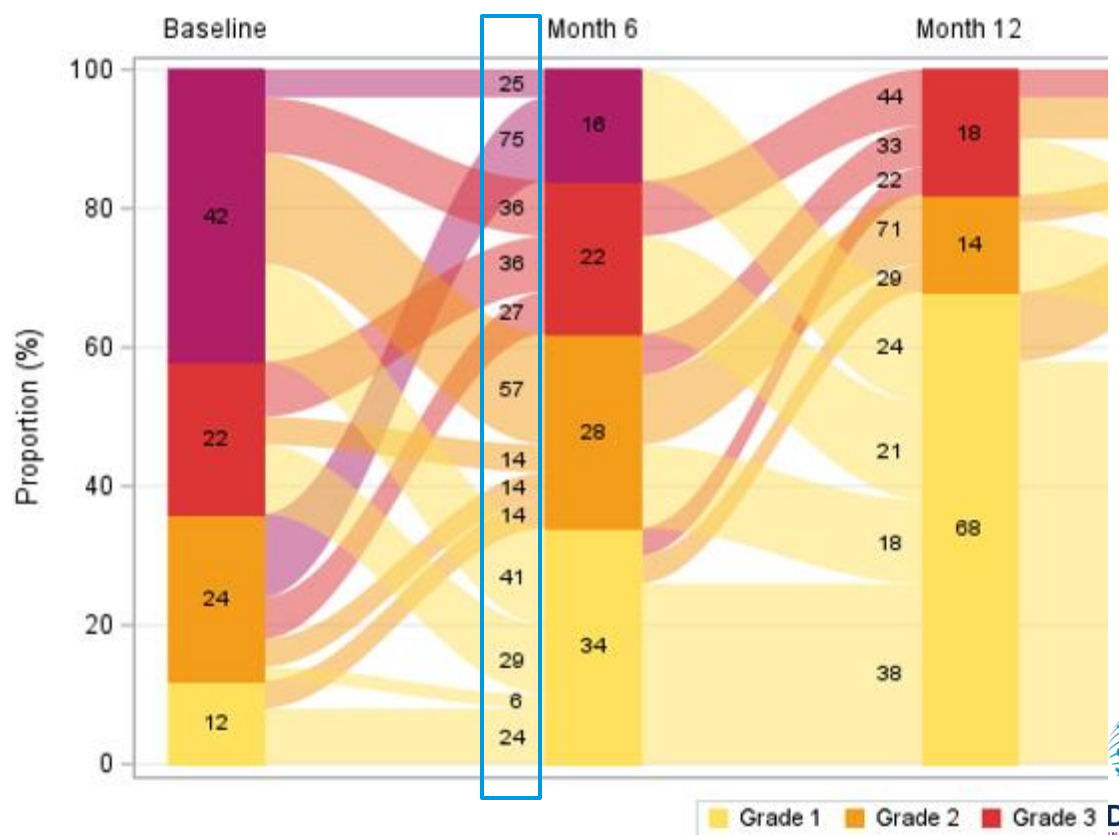
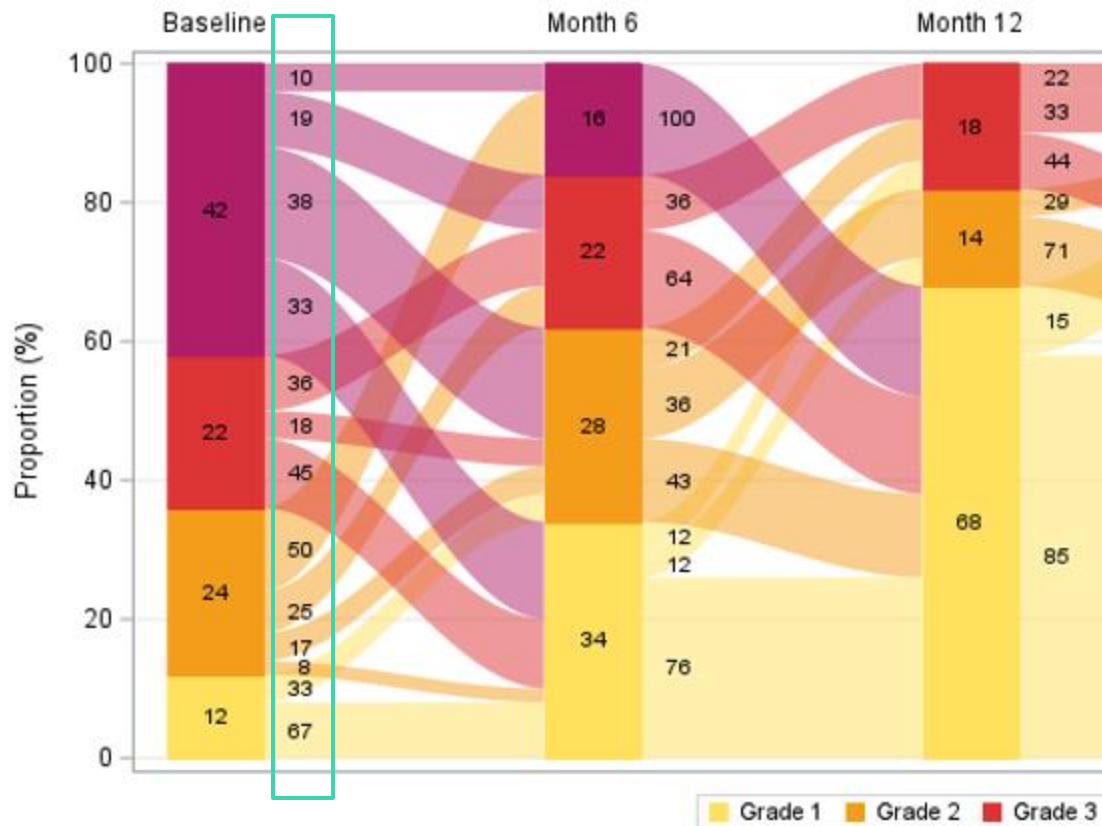
Color Customization

Swap inflow and outflow colors

Outflow color
and annotation

```
legendlabel dest=put(next_ybar,&yfmt.);  
%if &band color=INFLOW %then %do;  
band=tranwrd(band,'legendlabel','legendlabel dest');  
%end;
```

Inflow color and
annotation



DSSE
INJECTING THE DOTS

Other Improvements

Adjust how
many counts
or
percentages
are shown

Generate
missing
records based
on expected
visits

Unify the
separate
components
into single
macro

Agenda



Introduction

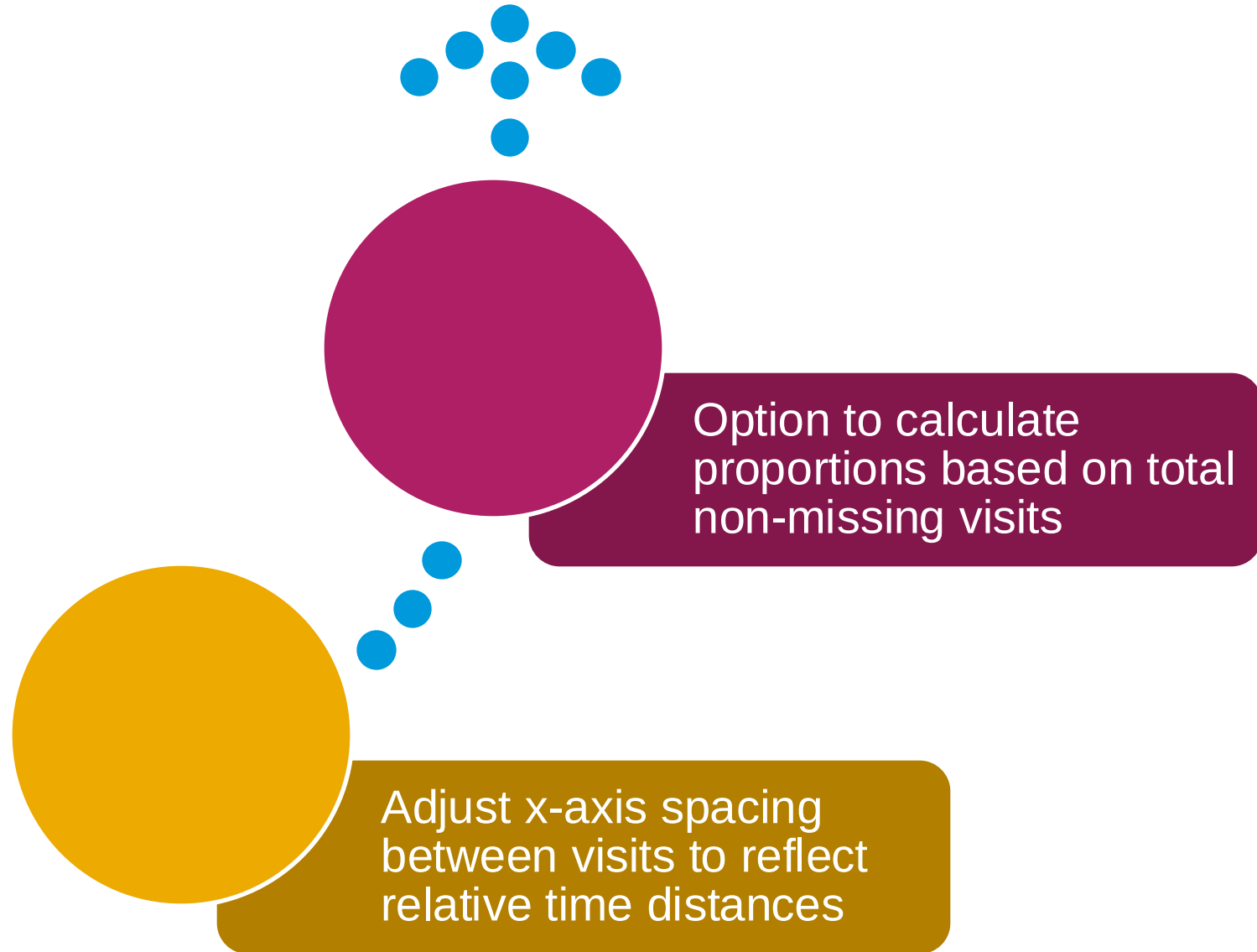
Motivation

Existing Solution

Improvement

Next Steps

Future Improvements



Conclusion

Visualize

Sankey plots are excellent visualizations to help understand patient journeys

Enhance

We updated the existing SAS macro to add improvements that assist in understanding the data

Analyze

This has applications to everyday analysis, particularly data normally presented in shift tables

Act

This allows for better decision-making and patient management

Q & A



| || Thank You!



DSSE
CONNECTING THE DOTS

 Alnylam®