

Visualizing PROC TRANSPOSE!

An Introduction

genzyme

How do I know I need to Transpose?

It'd sure be a lot easier if these observations were next to each other

- **The UP Transpose**

It'd sure be a lot easier if these variables were observations one underneath each other

- **The DOWN Transpose**

Four Questions About Your Transpose

1. What should stay the same? BY
2. What goes up? ID
3. What goes down? VAR
4. What goes into the middle? VAR

On Base Percentage

$$\text{OBP} = \frac{\text{H} + \text{BB} + \text{HBP}}{\text{AB} + \text{BB} + \text{HBP} + \text{SF}}$$

The Data – Calculate $H+BB+HBP/AB+BB+HBP+SF$

Team	Player	Stat	Value2007	Value2006
Orioles	Kevin Millar	H	121	117
Orioles	Kevin Millar	BB	76	59
Orioles	Kevin Millar	HBP	8	12
Orioles	Kevin Millar	AB	476	430
Orioles	Kevin Millar	SF	2	2
Orioles	Nick Markakis	H	191	143
Orioles	Nick Markakis	BB	61	43
Orioles	Nick Markakis	HBP	5	3
Orioles	Nick Markakis	AB	637	491
Orioles	Nick Markakis	SF	6	2

1. What should stay the same? The BY Statement

```
PROC TRANSPOSE DATA=Base_stats OUT=tran_stats;  
  BY team player;
```

1. What should stay the same? Visualization

Team	Player	Stat	Value2007	Value2006
Orioles	Kevin Millar	H	121	117
Orioles	Kevin Millar	BB	76	59
Orioles	Kevin Millar	HBP	8	12
Orioles	Kevin Millar	AB	476	430
Orioles	Kevin Millar	SF	2	2
Orioles	Nick Markakis	H	191	143
Orioles	Nick Markakis	BB	61	43
Orioles	Nick Markakis	HBP	5	3
Orioles	Nick Markakis	AB	637	491
Orioles	Nick Markakis	SF	6	2

2. What goes up? The ID Statement

Up Movement creates **VARIABLES** from unique **VALUES** of the ID variable

```
PROC TRANSPOSE DATA=Base_stats OUT=tran_stats;  
  BY team player;  
  ID stats;
```

2. What goes up? Visualization

Team	Player	Stat
Orioles	Kevin Millar	H
Orioles	Kevin Millar	BB
Orioles	Kevin Millar	HBP
Orioles	Kevin Millar	AB
Orioles	Kevin Millar	SF
Orioles	Nick Markakis	H
Orioles	Nick Markakis	BB
Orioles	Nick Markakis	HBP
Orioles	Nick Markakis	AB
Orioles	Nick Markakis	SF

3. What goes down? The VAR Statement

Individual variable **NAMES** are turned into **VALUES** of a new variable (**_NAME_**)

```
PROC TRANSPOSE DATA=Base_stats OUT=tran_stats;  
  BY team player;  
  ID stats;  
  VAR value2007 value2006;  
RUN;
```

3. What goes down? Visualization

Team	Player	_NAME_	H	BB	HBP	AB	SF
Orioles	Kevin Millar						
Orioles	Kevin Millar						
Orioles	Nick Markakis						
Orioles	Nick Markakis						

Value2007	Value2006
121	117
76	59
8	12
476	430
2	2
191	143
61	43
5	3
637	491
6	2

4. What goes into the middle? The VAR Statement

The **VALUES** of value2006, value2007 are pushed into the newly created variables (H,BB,HBP,AB,SF)

```
PROC TRANSPOSE DATA=Base_stats OUT=tran_stats;  
  BY team player;  
  ID stats;  
  VAR value2007 value2006;  
RUN;
```

4. What goes into the middle? Visualization

Team	Player	_NAME_	H	BB	HBP	AB	SF
Orioles	Kevin Millar						
Orioles	Kevin Millar						
Orioles	Nick Markakis						
Orioles	Nick Markakis						

Value2007	Value2006
121	117
76	59
8	12
476	430
2	2
191	143
61	43
5	3
637	491
6	2

Finished Product – Calculate H+BB+HBP/AB+BB+HBP+SF

Team	Player	_NAME_	H	BB	HBP	AB	SF	OBP
Orioles	Kevin Millar	Value2007	121	76	8	476	2	.365
Orioles	Kevin Millar	Value2006	117	59	12	430	2	.374
Orioles	Nick Markakis	Value2007	191	61	5	637	6	.362
Orioles	Nick Markakis	Value2006	143	43	3	491	2	.351

Questions??

Contact us at:

Daniel.Boisvert@Genzyme.com

www.DanielBoisvert.com

