

Wide vs. long data

Learning outcomes

By the end of this module, students will be able to:

- 1) Define long and wide data formats
- 2) List characteristics that differentiate long and wide data formats

wide data

| Student | Math | Science | English | Art |
|----------|------|---------|---------|-----|
| Samantha | 91 | 96 | 88 | 74 |
| Jacob | 98 | 94 | 75 | 95 |
| Zach | 81 | 89 | 93 | 99 |
| Bonnie | 78 | 74 | 82 | 82 |

long data

| Student | Subject | Grade |
|----------|---------|-------|
| Samantha | Math | 91 |
| Samantha | Science | 96 |
| Samantha | English | 88 |
| Samantha | Art | 74 |
| Jacob | Math | 98 |
| Jacob | Science | 94 |
| Jacob | English | 75 |
| Jacob | Art | 95 |
| Zach | Math | 81 |
| Zach | Science | 89 |
| Zach | English | 93 |
| Zach | Art | 99 |
| Bonnie | Math | 78 |
| Bonnie | Science | 74 |
| Bonnie | English | 82 |
| Bonnie | Art | 82 |

wide vs. long

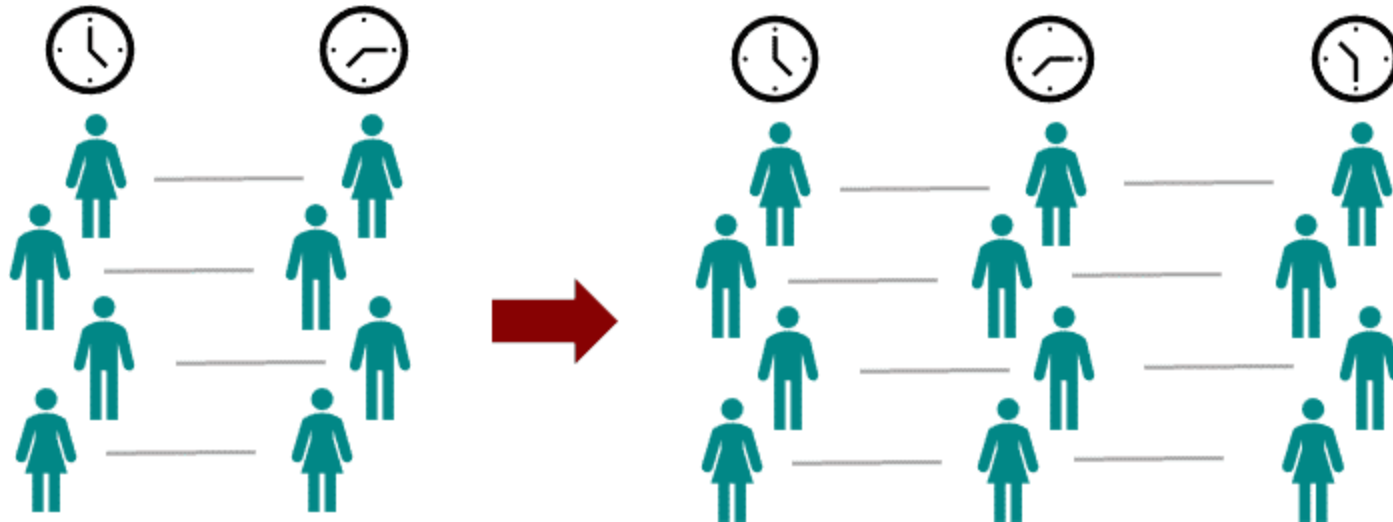


| Student | Math | Science | English | Art |
|----------|------|---------|---------|-----|
| Samantha | 91 | 96 | 88 | 74 |
| Jacob | 98 | 94 | 75 | 95 |
| Zach | 81 | 89 | 93 | 99 |
| Bonnie | 78 | 74 | 82 | 82 |

| Student | Subject | Grade |
|----------|---------|-------|
| Samantha | Math | 91 |
| Samantha | Science | 96 |
| Samantha | English | 88 |
| Samantha | Art | 74 |
| Jacob | Math | 98 |
| Jacob | Science | 94 |
| Jacob | English | 75 |
| Jacob | Art | 95 |
| Zach | Math | 81 |
| Zach | Science | 89 |
| Zach | English | 93 |
| Zach | Art | 99 |
| Bonnie | Math | 78 |
| Bonnie | Science | 74 |
| Bonnie | English | 82 |
| Bonnie | Art | 82 |

real world examples

| Subject | Resting Heart Rate 1 Month Before Training Program | Resting Heart Rate in Middle of Training Program | Resting Heart Rate 1 Month After Training Program |
|----------|--|--|---|
| Michael | 65 | 58 | 60 |
| Dwight | 55 | 48 | 49 |
| Andy | 58 | 55 | 55 |
| Meredith | 68 | 60 | 64 |
| Angela | 47 | 45 | 45 |



2001

| Year | State | Sales revenue |
|------|----------------|---------------|
| 2001 | California | 1704210.8 |
| 2001 | Colorado | 448301.5 |
| 2001 | DC | 693210.5 |
| 2001 | Florida | 405985.1 |
| 2001 | Illinois | 738223.8 |
| 2001 | Massachussetts | 238818.7 |
| 2001 | New York | 1667695.8 |
| 2001 | Texas | 2199677.4 |

2002

| Year | State | Sales revenue |
|------|----------------|---------------|
| 2002 | California | 2782679.5 |
| 2002 | Colorado | 768389.5 |
| 2002 | DC | 1215158 |
| 2002 | Florida | 661249.8 |
| 2002 | Illinois | 1150658.8 |
| 2002 | Massachussetts | 157718.7 |
| 2002 | New York | 2763503.1 |
| 2002 | Texas | 3732888.6 |

Summary

- Wide data spreads measurements across columns
- Long data stacks observations into rows
- Long format works better with many Tidyverse packages