

# Quantitative Variables: Stemplots

## Lecture 4

Robb T. Koether

Hampden-Sydney College

Mon, Jan 18, 2016

# Outline

- 1 Interpreting the Graph
- 2 Stemplots
- 3 Splitting Stems
- 4 Back-to-back Stemplots
- 5 Assignment

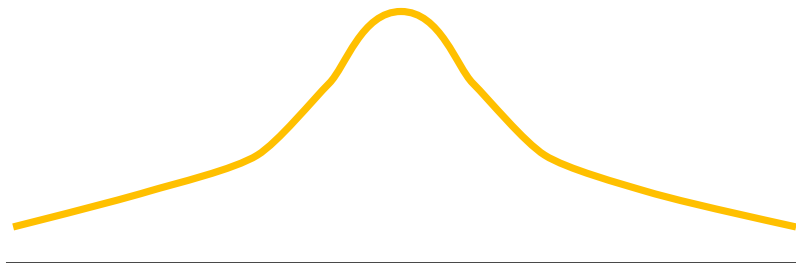
# Outline

- 1 Interpreting the Graph
- 2 Stemplots
- 3 Splitting Stems
- 4 Back-to-back Stemplots
- 5 Assignment

# Interpreting the Graph

- In a display of quantitative data, what do we wish to convey?

# Interpreting the Graph



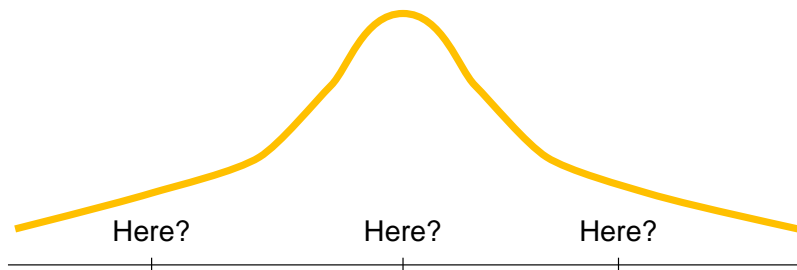
Are the data clustered around the middle?

# Interpreting the Graph



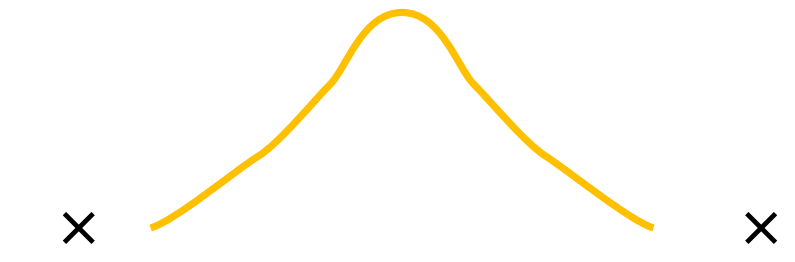
Or are they spread out?

# Interpreting the Graph



Where is the middle?

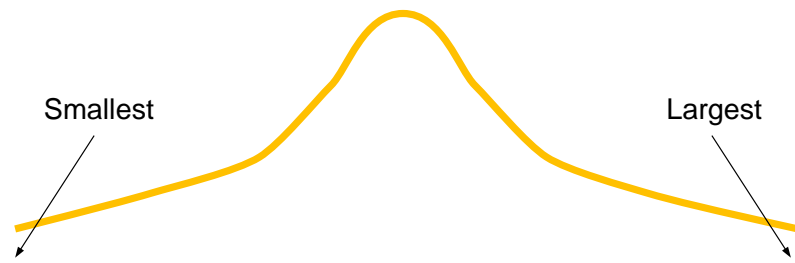
# Interpreting the Graph



Are they any outliers?

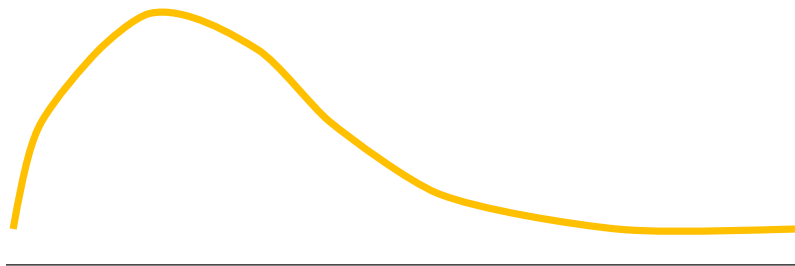


# Interpreting the Graph



What are the extremes?

# Interpreting the Graph



Are the values skewed to the left or right?

# Interpreting the Graph

- Shape – Does the shape have a single peak, more than one peak, or no peak?
  - **Symmetric** – The left side is a mirror image of the right side.
  - **Skewed left** – single peak, long tail to the left.
  - **Skewed right** – single peak, long tail to the right.
- **Center** – What appears to be the center (or balancing point) of the distribution?
- **Spread** – Are the data tightly clustered or are they spread widely?
- Are there any **outliers**? (What constitutes an outlier?)

# Outline

- 1 Interpreting the Graph
- 2 Stemplots**
- 3 Splitting Stems
- 4 Back-to-back Stemplots
- 5 Assignment

# Stemplots

- A very simple, and often satisfactory, way to draw a distribution is to create a **stemplot**.
- To do this, we split each number (datum) into two parts, the left part (stem), and the digit to the right (leaf).

# Stemplots

- For example, the value 12.34 could be split as
  - 1|234 – stem = 1, leaf = 2, or
  - 12|34 – stem = 12, leaf = 3, or
  - 123|4 – stem = 123, leaf = 4.
- Indeed, it could even be split as
  - |1234 – stem = 0, leaf = 1, or
  - 1234| – stem = 1234, leaf = 0.
- Note that the decimal point is irrelevant.

# Stemplots

## Definition (Stem and Leaf)

Given a number, the first digit after the split is that number's **leaf**. The digits to the left of the split form its **stem**.

## Definition (Stemplot)

A **stemplot** of a quantitative variable is a display which is arranged in two columns. The left column shows the full range of stems that span the data. For every stem in the left column, the right column shows the leaf of every value that had that stem.

## Example (Rainfall Data)

- Draw a stemplot of rainfall data for August in Richmond, VA (1986 - 2015).

6.74	1.24	4.04	4.90	5.72	2.88
6.91	5.58	2.52	8.42	4.44	1.41
1.84	2.00	2.79	2.30	3.15	3.59
16.02	2.56	5.99	6.81	5.73	4.04
3.92	7.10	3.50	7.64	3.61	2.77

(16.02 inches occurred in 2004.)

- Draw a stemplot.



# Outline

- 1 Interpreting the Graph
- 2 Stemplots
- 3 Splitting Stems**
- 4 Back-to-back Stemplots
- 5 Assignment

## Example (Rainfall Data)

- The rainfall data in inches.

6.74	1.24	4.04	4.90	5.72	2.88
6.91	5.58	2.52	8.42	4.44	1.41
1.84	2.00	2.79	2.30	3.15	3.59
16.02	2.56	5.99	6.81	5.73	4.04
3.92	7.10	3.50	7.64	3.61	2.77

(16.02 inches occurred in 2004.)

## Example (Rainfall Data)

- The same data expressed in centimeters.

17.12	3.15	10.26	12.45	14.53	7.32
17.55	14.17	6.40	21.39	11.28	3.58
4.67	5.08	7.09	5.84	8.00	9.12
40.69	6.50	15.21	17.30	14.55	10.26
9.96	18.03	8.89	19.41	9.17	7.04

(40.69 cm occurred in 2004.)

# Example

## Example (Percent Born Outside US)

State	Perc	State	Perc	State	Perc	State	Perc
AL	2.8	IN	4.2	NE	5.6	SC	4.1
AK	7.0	IA	3.8	NV	19.1	SD	2.2
AZ	15.1	KS	6.3	NH	5.4	TN	3.9
AR	3.8	KY	2.7	NJ	20.1	TX	15.9
CA	27.2	LA	2.9	NM	10.1	UT	8.3
CO	10.3	ME	3.2	NY	21.6	VT	3.9
CT	12.9	MD	12.2	NC	6.9	VA	10.1
DE	8.1	MA	14.1	ND	2.1	WA	12.4
FL	18.9	MI	5.9	OH	3.6	WV	1.2
GA	9.2	MN	6.6	OK	4.9	WI	4.4
HI	16.3	MS	1.8	OR	9.7	WY	2.7
ID	5.6	MO	3.3	PA	5.1		
IL	13.8	MT	1.9	RI	12.6		

# Outline

- 1 Interpreting the Graph
- 2 Stemplots
- 3 Splitting Stems
- 4 Back-to-back Stemplots**
- 5 Assignment

# Back-to-back Stemplots

- Sometimes we want a display that will compare two distributions to each other.
- We will compare rainfall in August (a dry month?) with rainfall in April (a wet month?).

# Back-to-back Stemplots

## Definition

A **back-to-back stemplot** contains two stemplots in one display. The leaves of one stemplot are on the left and the leaves of the other stemplot are on the right. A set of common stems is in the middle.

Leaves	Stems	Leaves
...	⋮	...
...	⋮	...
...	⋮	...

# Example

## Example (Back-to-back Stemplots)

- Rainfall data for April, Richmond, VA (1986 - 2015).

1.08	6.67	2.05	4.17	2.73	0.80
1.93	4.13	2.70	11.12	4.85	3.94
2.79	2.57	4.12	2.14	2.33	4.37
3.42	2.03	2.18	3.62	8.32	2.56
1.59	2.63	2.40	2.88	4.05	5.33



# Back-to-back Histograms

- See the Census Bureau's display for the male and female age distributions: [Click here](#)

# Outline

- 1 Interpreting the Graph
- 2 Stemplots
- 3 Splitting Stems
- 4 Back-to-back Stemplots
- 5 Assignment**

# Assignment

## Homework

- Read Chapter 1: Quantitative Variables: Stemplots.
- Apply Your Knowledge: 1.10.
- Check Your Skills: 1.19, 1.20, 1.21.
- Exercises: 36, 37, 38.