✢ Beginner user

Fit Y-by-X a Tutorial

Karen Copeland, Ph.D.

Discovery 2014

Boulder Statistics

www.boulderstats.com

Outline of Tutorial

- Start with data from the Boulder flood to:
  - Import data from text file
  - Explore the Fit Y-by-X platform (Bivariate)
- Generate fictitious (simulated) data to:
  - Continue to explore the Fit Y-by X platform (One-Way, Logistic, Contingency)
- Along the way we will explore JMP gems
  - Table Features: filters, summarize
  - Column Features: properties, initialization, formulas, labels, viewer
  - Row Features: color/mark by, next selected, matching, labels
  - Interactive Features: adding columns from selected data, copying axis settings, copying frame contents, customizing platforms

Live Demo….

Daily Precip from 1887 for Boulder Co

• File>Open>Boulder Rain from 1887.txt (Open As: Data using preview)
Open Data Table From Text
• Use "right click>select matching cell>control x" to remove -998 (from precip column)
• plot precip by mon
precip by mon
• set auto recalc (red triangle>script>automatic recalc)
• highlight a problem value
• use next slected (Rows>Next Selected)
• remove matching cells (control x)
• hover over point
• make year a label (Columns>Label)
• use select to see # of points
• pin
• swap x and y
• add local data filter (red triangle>script>local data filter) and explore amounts and years.
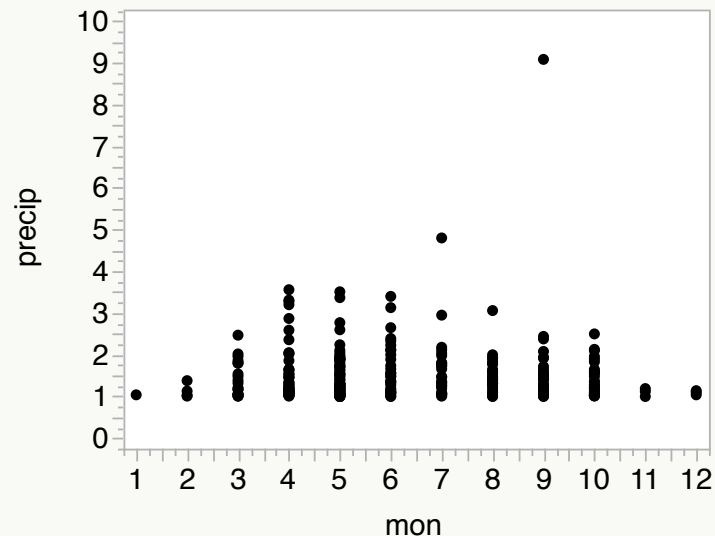
Local Data Filter

297 matching rows
☐ Inverse

1914 ≤ year ≤		2014

1 ≤ precip ≤		9.08

Bivariate Fit of precip By mon

8 Days of Rain Formula

- Add a new column (menu or double click to right of last column)
- Add a new formula (menu or right click on column heading)
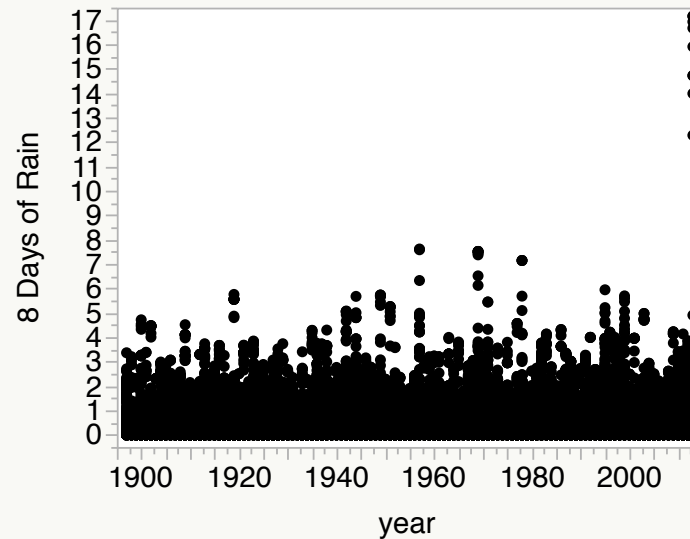- Summation(i = Row() - 8, Row(), :precip[i, Empty()])

Local Data Filter

42974 matching rows

☐ Inverse

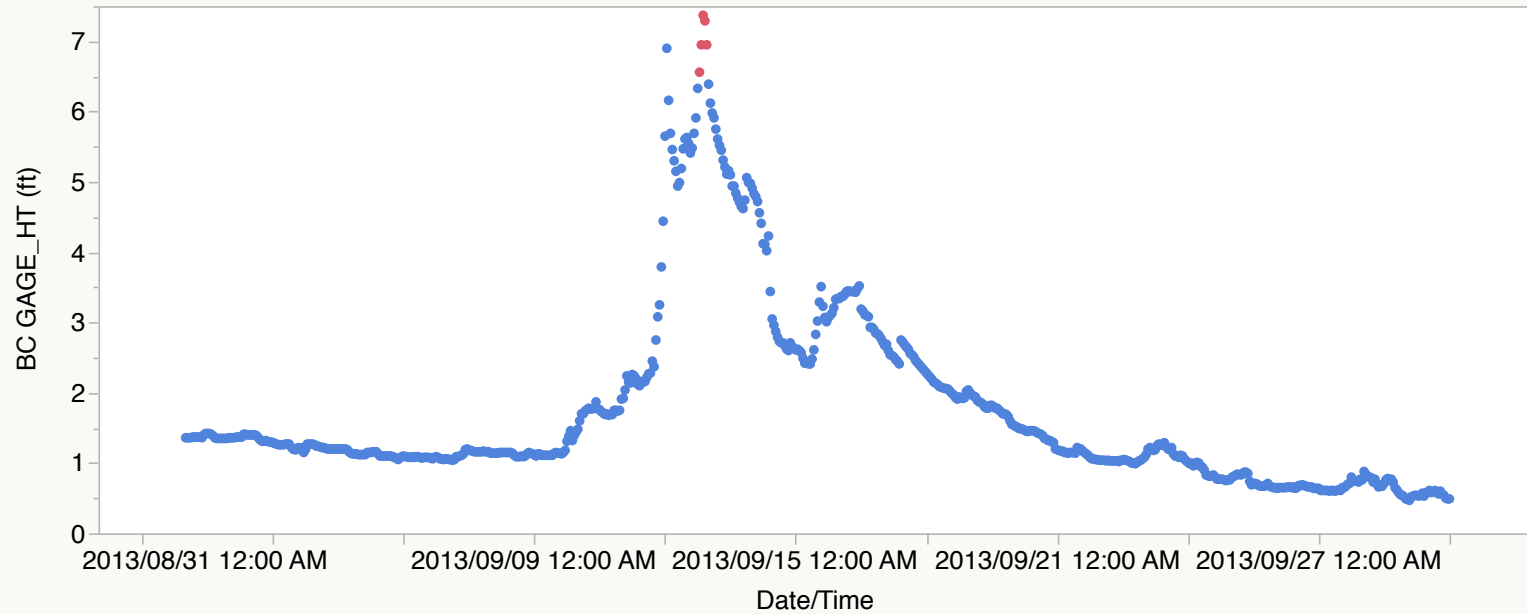| 0 ≤ 8 Days of Rain ≤ | 17.15 |

Bivariate Fit of 8 Days of Rain By year

SV and BC Hourly Sept 2013

- Ht by Date/Time
- Select and then right click>Name Selection in Column
- Row>Color or Mark by Column
- Show Column Viewer

Bivariate Fit of BC GAGE_HT (ft) By Date/Time

Name Selection in Column...

Label the currently selected rows and save the value(label) in a column.

Column Name	
Selected	1
Unselected	0

Cancel   OK

Build a Simulated Diagnostic Data Set

- File>New Table
- Rows>Add Rows...100
- Column Info (Diagnosis: Character) >Initialize Data (Sequence Data: Positive, Negative)
- Add Column (SID: numeric)>Initialize Data (sequence)
- Add Column (Result: numeric)>Formula
- Formula: Match(:Diagnosis,
    "Negative", Random Lognormal(1.3, 0.5),
    "Positive", Random Normal() * 2 + 8,
    Empty(), Empty()
)
- Add Column (Test Outcome: character)>Formula
- Formula: If(:Result < 4,
    "Neg",
    "Pos"
)
- Add Column (Group: numeric)> Initialize Data (random indicator)
- Column Properties > Character
- Columns>Recode

More Fit Y by X

Contingency Analysis
- Nominal by Nominal = Mosaic

Logistic Regression
- Nominal by Continuous = Logistic
- Right Click > Customize...
- Red Triangle>ROC Curve
- Right Click > Customize...use script
- Right Click> Copy Frame Contents

Oneway Analysis
- Continous by Nominal = Oneway
- t-tests
- ANOM (Analysis of Means)
- Other multiple comparison tests
- Summary stats