trelliscope

D&R visualization tool
http://tessera.io

Barret Schloerke
Purdue University
Bill Cleveland & Ryan Hafen

Motivation

- 'Blind' analysis can be very dangerous

  - Miss key data structure

  - Visualization is a critical component

- Visualization handles **very** complex structures

  - Multiple variables

  - Multiple types of variables (continuous, discrete, date, location, etc.)
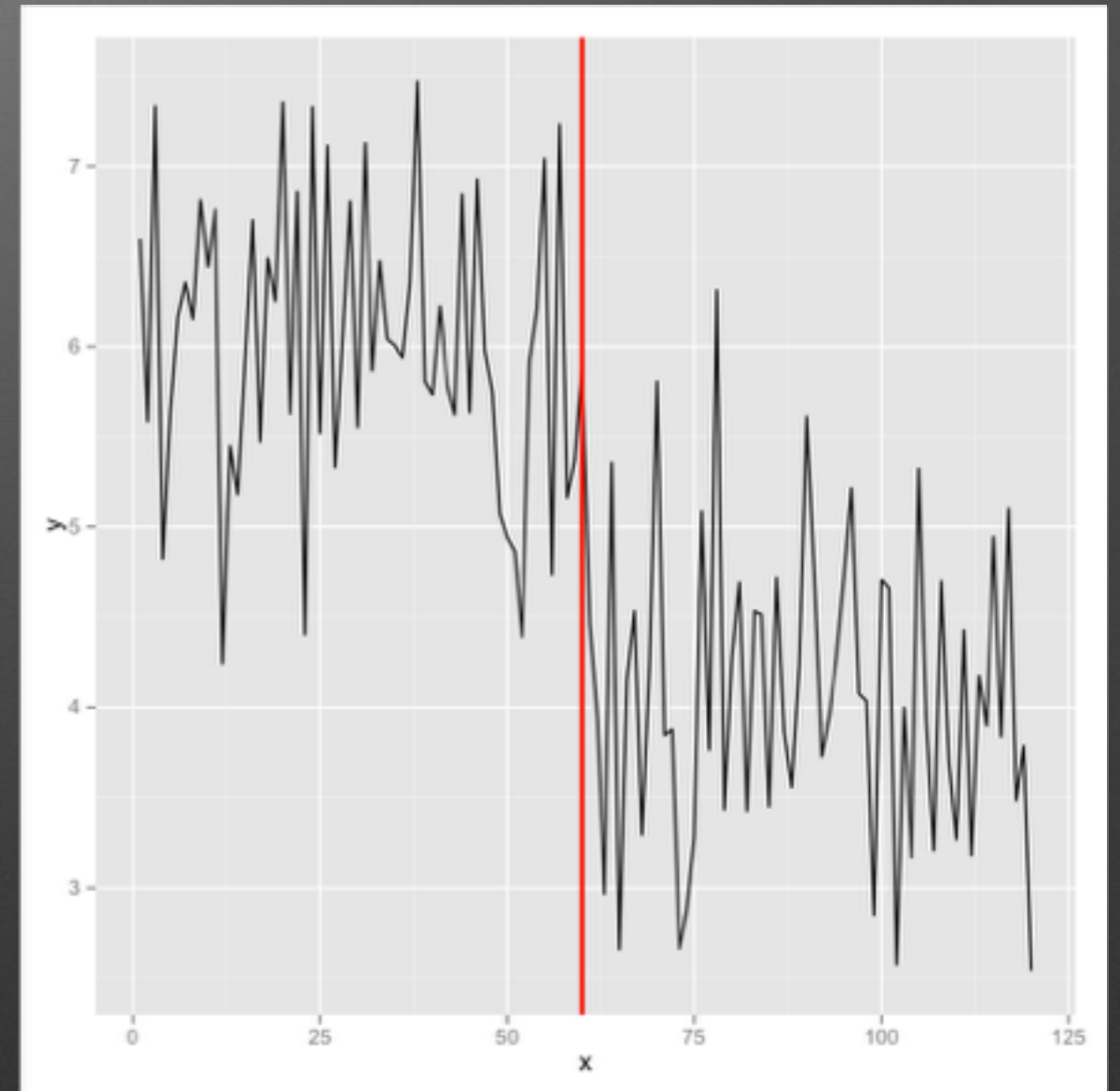
Example

- Financial time series data
  - collected every milli-sec
  - 1 TB / day

- Can not display **all** data in one picture
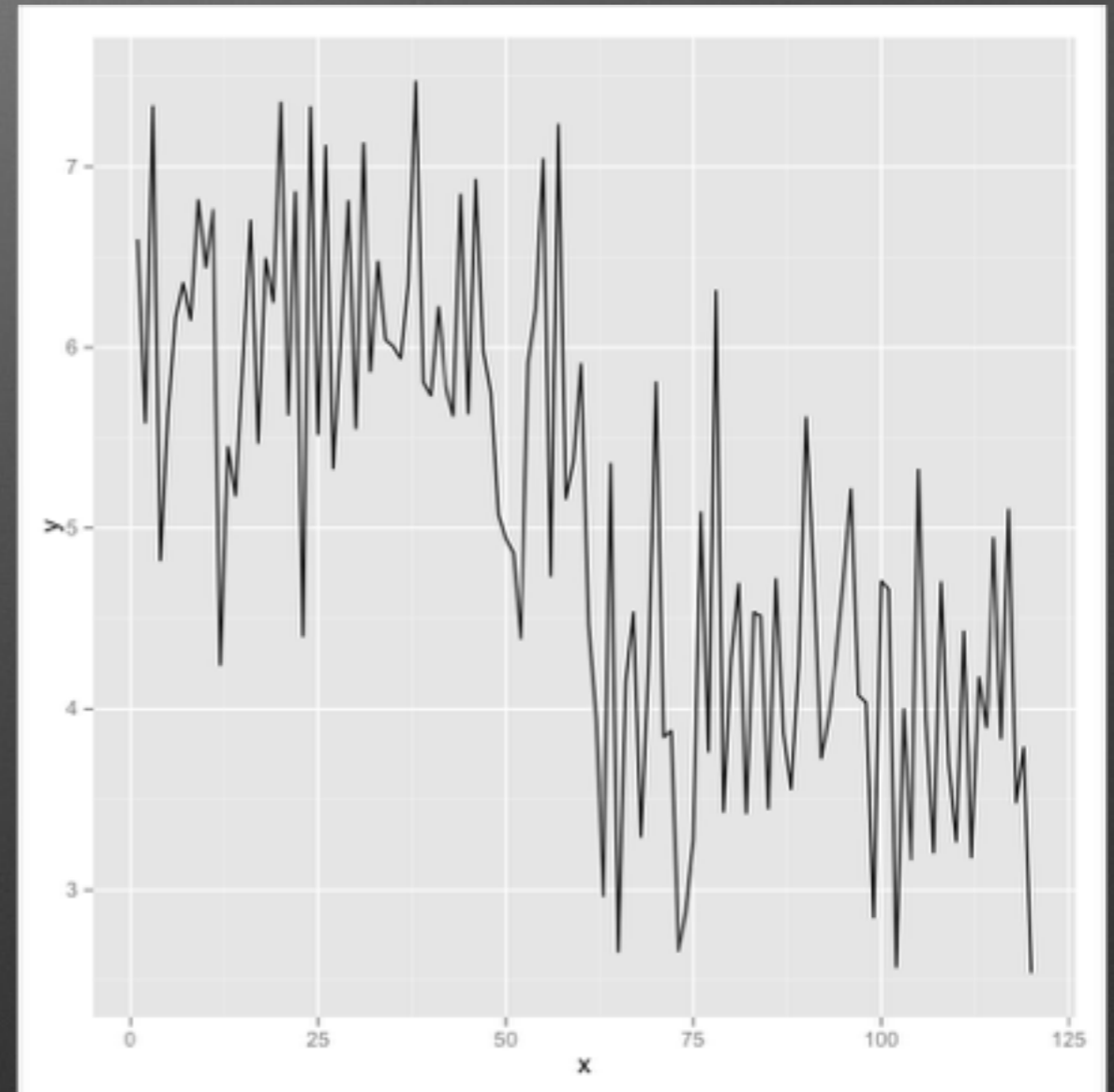
Example

- Financial time series data
  - collected every milli-sec
  - 1 TB / day

- Can not display **all** data in one picture

- Need to detect
  - Change points
  - Outliers

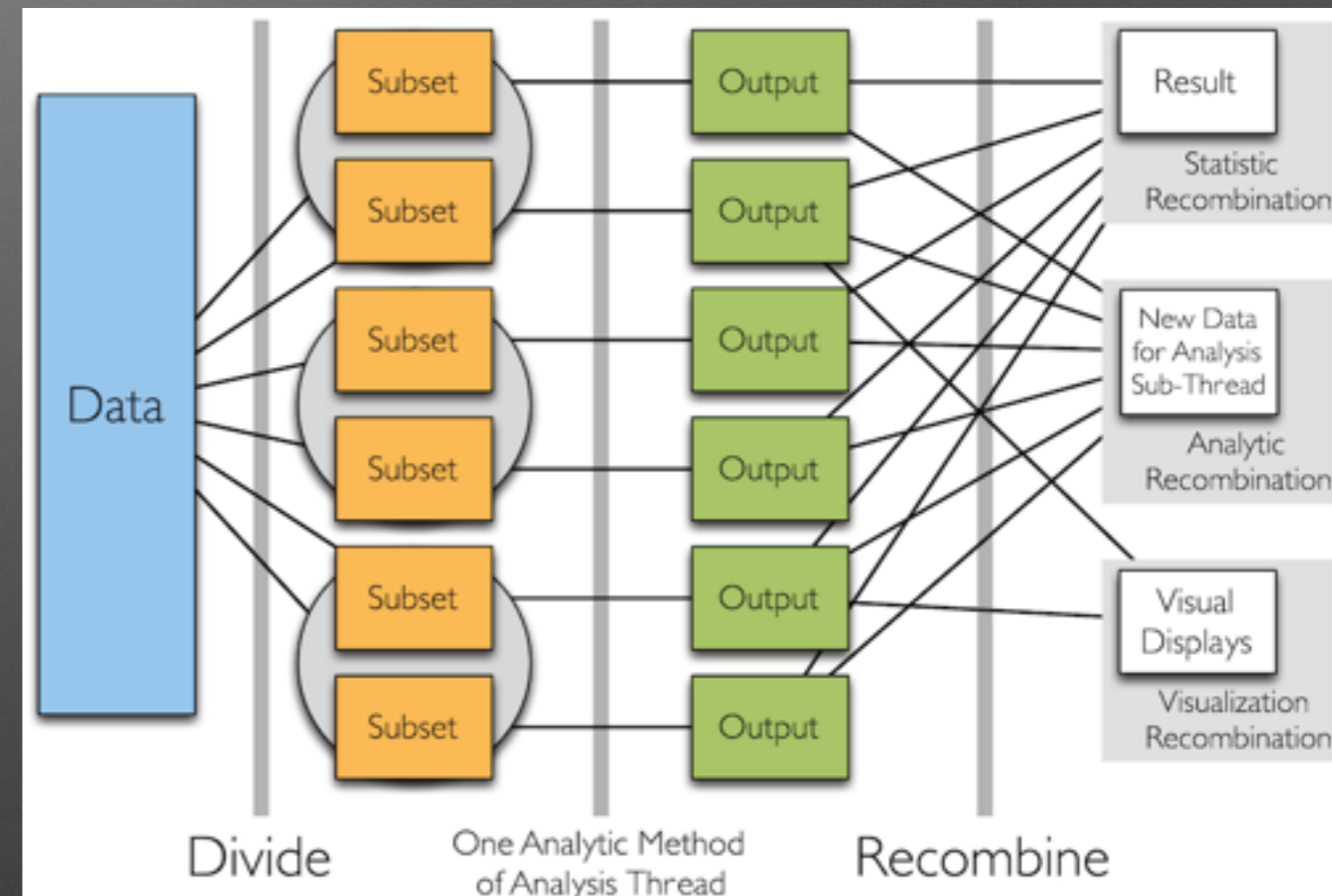- Trelliscope!

Trelliscope

- http://tessera.io/docs-trelliscope/

- Divide & Recombine (D&R) visualization tool

- Based on Trellis Display

  - Effective approach to visualizing complex data

- Enables scalable, flexible, detailed visualization of data
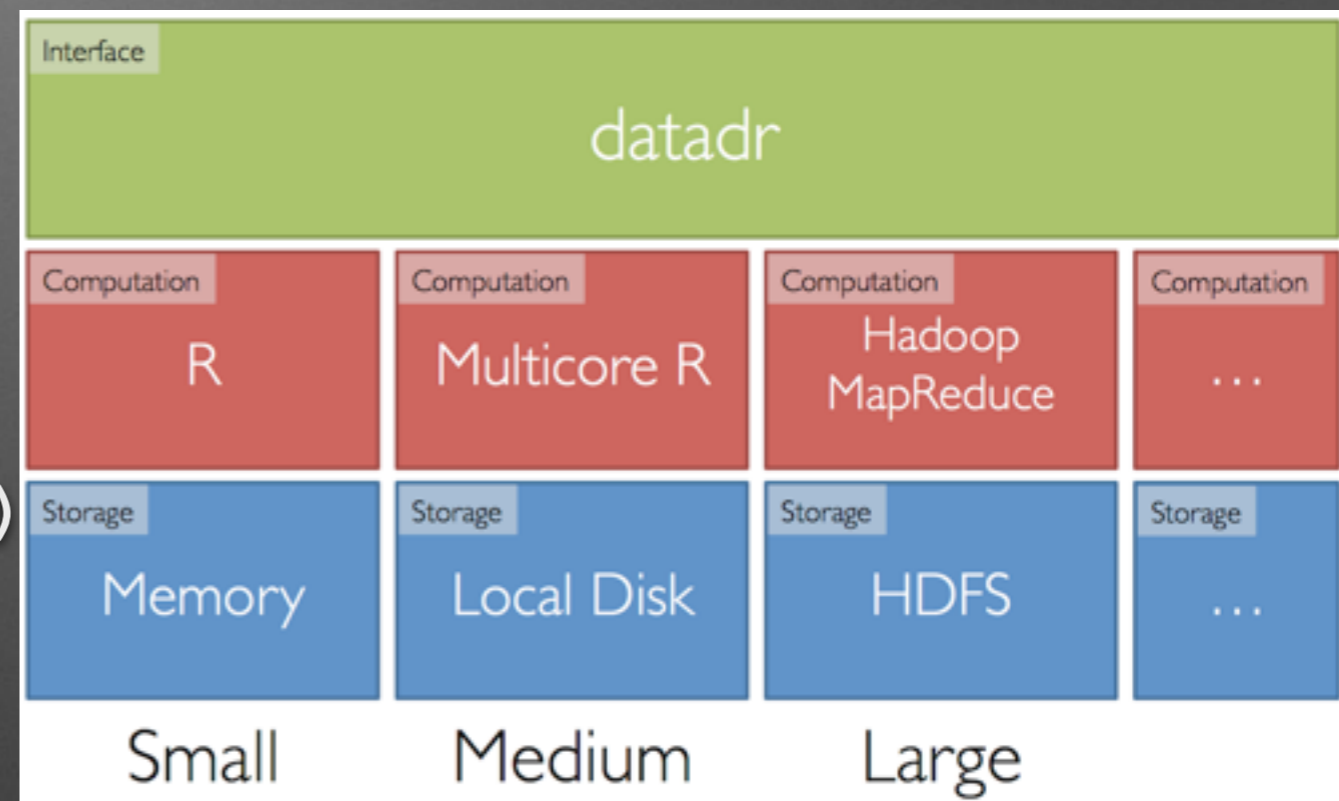
http://tessera.io Stack

- Divide and Recombine

http://tessera.io Stack

- Divide and Recombine

- datadr

  - Interface for D&R

  - Back-end agnostic

    - RHIPE (Hadoop MapReduce)

    - Spark

    - Local Disk

- Coding only within R

http://tessera.io Stack

- Divide and Recombine

- datadr
  - Interface for D&R
    - Back-end agnostic
      - RHIPE (Hadoop MapReduce)
      - Spark
      - Local Disk
  - Coding only within R

- Trelliscope
  - Visualization of D&R data

Functionality

- Apply functions
  - Visual display
    - panel function
  - Cognostics
    - computational diagnostics
    (metrics)

- Interactively
  - Sort
  - Filter

- Change
  - Page number
  - Panel layout
  - Labels displayed

Trelliscope

View Options

- Panel Layout
- Panel Function
- Panel Labels
- Related Displays

Cognostics

- Active Cognostics
- Table Sort / Filter
- Univariate Filter
- Bivariate Filter
- Multivariate Filter
- Sample

Open a New Display

Display	Thumb
fields / **list_vs_time_ggplot** : List and sold priceover time w/ggplot2 *2854 panels, pre-rendered=false, data storage=kvMemory, updated: 2015-02-24 00:09:02*	
fields / **list_vs_time_ex1** : Example 1 *2854 panels, pre-rendered=false, data storage=kvMemory, updated: 2015-02-24 00:09:49*	
fields / **list_vs_time_ex2** : Example 2 *2854 panels, pre-rendered=false, data storage=kvMemory, updated: 2015-02-24 00:10:09*	

Close

Trelliscope

fields / list_vs_time_ggplot

1 / 2854 ← → 1x ▾

View Options

▦ Panel Layout

✐ Panel Function

▤ Panel Labels

⊕ Related Displays

Cognostics

✹ Active Cognostics

▮ Table Sort / Filter

▮ Univariate Filter

∿ Bivariate Filter

∿ Multivariate Filter

⣿ Sample

county	Abbeville County
state	SC

Trelliscope

View Options

- ⊞ Panel Layout
- ✎ Panel Function
- ☰ Panel Labels
- ⊞ Related Displays

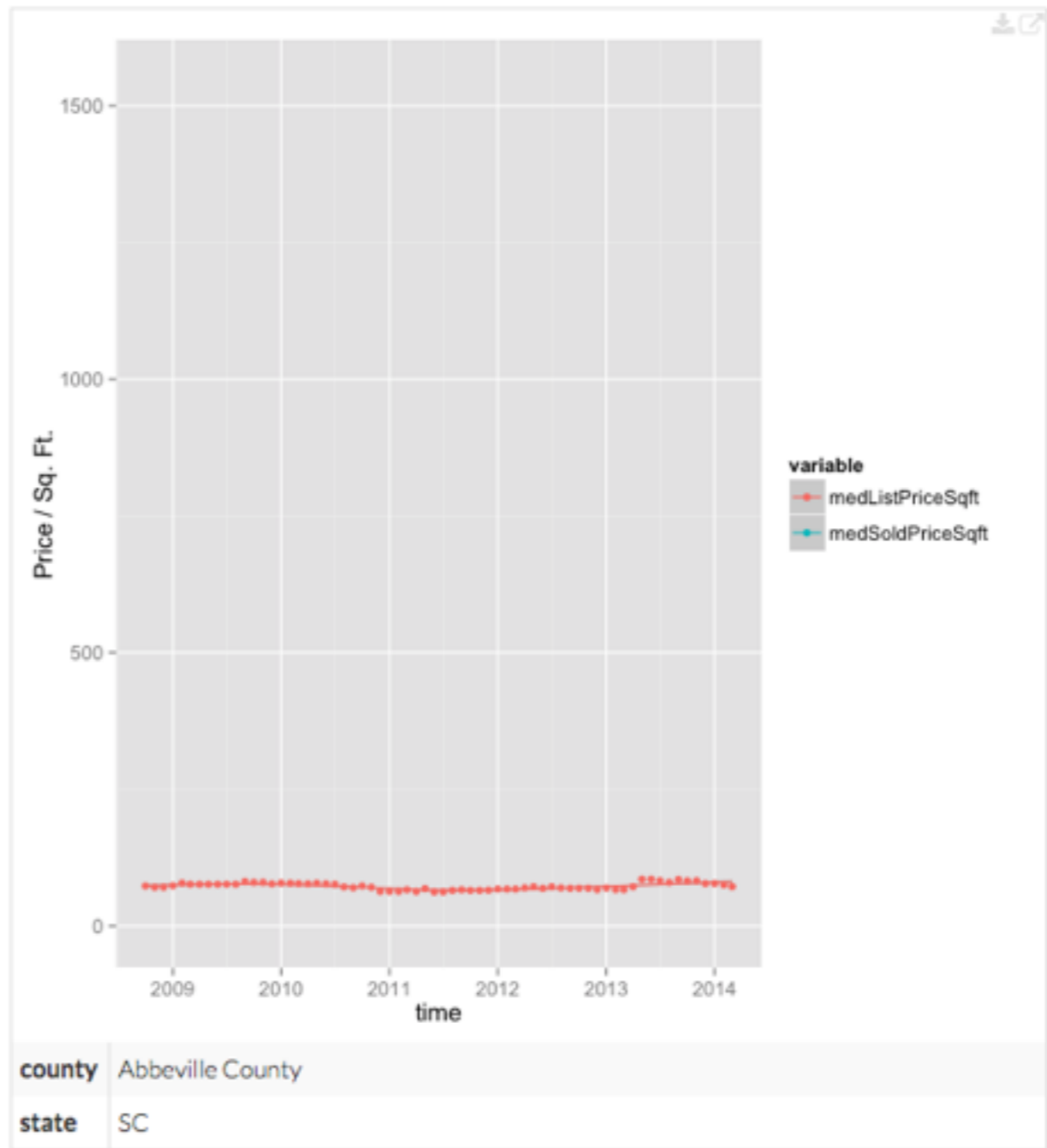**Cognostics**

- ✹ Active Cognostics
- ⬇ Table Sort / Filter
- ⥧ Univariate Filter
- ⤳ Bivariate Filter
- ⤢ Multivariate Filter
- ⣿ Sample

Cognostics View / Sort / Filter

View cognostics in a table and specify sort order or filtering of panels.
Shift-click on the panel header sorting buttons for multi-column sorting.

county ⇕	state ⇕	slope ⬇	meanList ⇕	meanSold ⇕	nObs ⇕
New York County	NY	0.24381422	1,113.30647	1,089.42031	6
Harney County	OR	0.16029776	113.17146	NaN	1
Kossuth County	IA	0.10273192	54.18891	NaN	
Santa Barbara County	CA	0.08631320	320.35298	236.23978	6
District of Columbia	DC	0.08214633	369.05464	422.47440	6
Knox County	NE	0.08081005	88.57545	33.97869	
Howard County	IA	0.07111307	54.99439	NaN	
Mills County	TX	0.06996998	182.08283	NaN	1
Texas County	OK	0.06808895	62.11879	56.38022	
St. Helena Parish	LA	0.06261914	75.69734	NaN	
regex	regex	0	From	From	From
	select ▲	To	To	To	To

Showing entries 1 - 10 of 1133

|◀ ◀ 1 of 114 ▶ ▶|

Columns to show

- panelKey
- county
- state
- slope ▼⬇
- meanList
- meanSold
- nObs
- zillowHref

Cancel   Apply

127.0.0.1:8100/#name=list_vs_time_ggplot&group=fields&sort=slope:desc&filter=slope:(from:0)&labels=county,state

Trelliscope

View Options

- ▦ Panel Layout
- ✎ Panel Function
- ▤ Panel Labels
- ⊕ Related Displays

Cognostics

- ✿ Active Cognostics
- ▥ Table Sort / Filter
- ▮▮ Univariate Filter
- ∿ Bivariate Filter
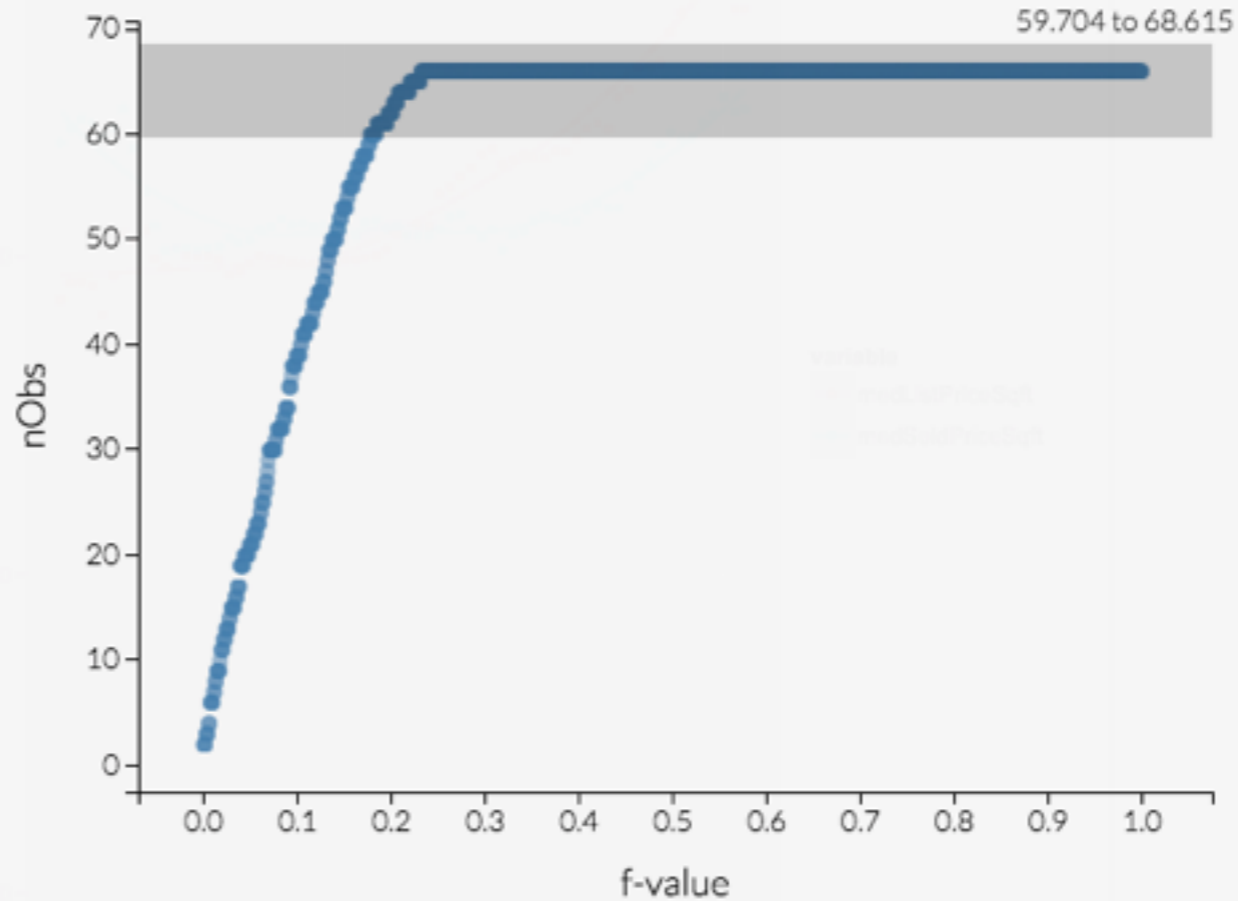- ⬚ Multivariate Filter
- ⠿ Sample

▮▮ Visual Univariate Filter

Filter panels based on a single cognostic guided by a visual display of the distribution of the cognostic.
Select a variable to view and click and drag on the display to specify filtering.

Variable to plot

slope	▼
meanList	
meanSold	
nObs	▼

59.704 to 68.615

[scatter plot with y-axis labeled "nObs" ranging 0 to 70, x-axis labeled "f-value" ranging 0.0 to 1.0]

[**Quantile** | Histogram | Bar]   [**Filtered** | All]

Cancel   **Apply**

State: ▼ slope ✕ ▥ slope ✕

Trelliscope Viewer ×

127.0.0.1:8100/#name=list_vs_time_ggplot&group=fields&sort=slope:desc&filter=slope:(from:0),nObs:(from:59.7035416216216;to:68.6151956756757)&labels=co...

Barret

Trelliscope

Panel Function Editor
Edit the code that generates each panel.

```
1   panelFn <-
2   structure(function(x) {
3     x %>%
4       select(time, medListPriceSqft, medSoldPriceSqft) %>%
5       gather(key = "variable", value = "value", medListPriceSqft, medSoldPriceSqft, -time) %>%
6       ggplot(aes(x = time, y = value, color = variable)) +
7         geom_smooth() +
8         geom_point() +
9         ylim(yRanges) +
10        labs(y = "Price / Sq. Ft.")
11  }, class = c("function", "ggplotFn"))
```

Note: this feature is currently disabled - panels will not update when you click 'Apply'...

Cancel   Apply

View Options

- ▦ Panel Layout
- ✍ Panel Function
- ▤ Panel Labels
- ⊕ Related Displays

Cognostics

- ❋ Active Cognostics
- ⬇ Table Sort / Filter
- ▥ Univariate Filter
- ⋏ Bivariate Filter
- ⬕ Multivariate Filter
- ⣿ Sample

State: ▼ slope × ▼ nObs × ⬍ slope ×

Trelliscope Viewer ×

Barret

127.0.0.1:8100/#name=list_vs_time_ggplot&group=fields&sort=slope:desc&filter=slope:(from:0),nObs:(from:59.7035416216216;to:68.6151956756757)&labels=co...

Trelliscope
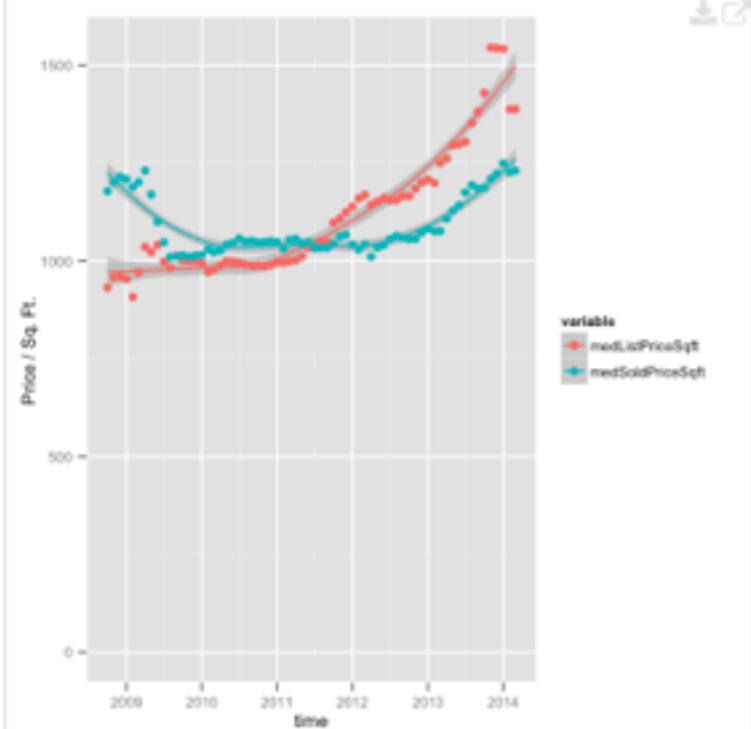
View Options

▦ Panel Layout

✎ Panel Function

▤ Panel Labels

⊕ Related Displays

Cognostics

❀ Active Cognostics

⫼ Table Sort / Filter

⏶ Univariate Filter

⋀ Bivariate Filter

⋁ Multivariate Filter

⋮ Sample

▦ Panel Layout

Specify the arrangement of the panels on each page by entering the number of rows and columns of panels per page. Future functionality for fixing / varying levels of cognostics per page to come.   ✕

Rows

1

Columns

3

Fixed Per Page

▼

Varies Per Page

▼

Panel Arrangement

| By Row | By Column |

Cancel   Apply

Trelliscope Viewer   ×

127.0.0.1:8100/#name=list_vs_time_ggplot&group=fields&sort=slope:desc&filter=slope:(from:0),nObs:(from:59.7035416216216;to:68.6151956756757)&labels=co...

Trelliscope

Panel Labels

Specify which cognostics to view underneath each panel.
Select cognostics to view by clicking and/or dragging the 'Show' column.

View Options
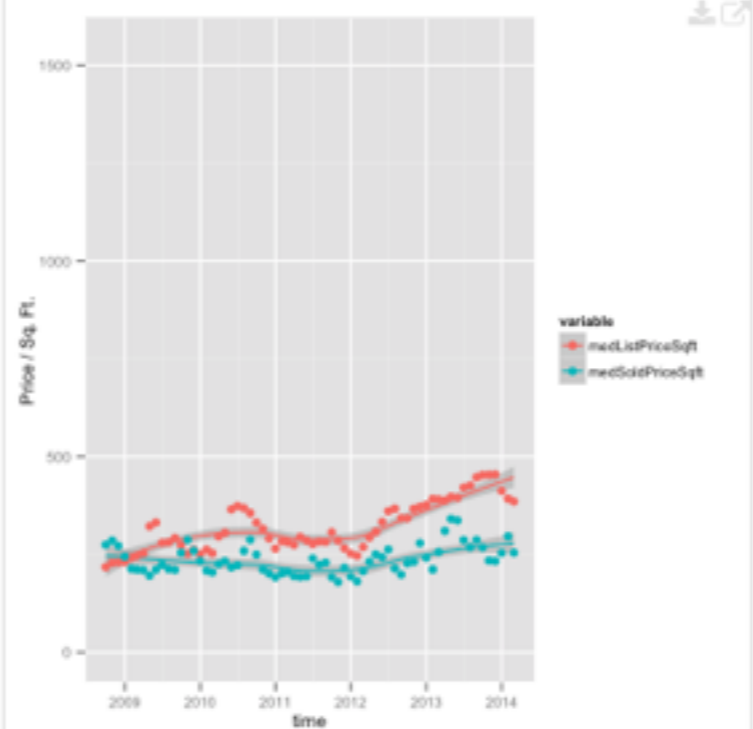
⊞ Panel Layout

✐ Panel Function

▤ Panel Labels

⊕ Related Displays

Cognostics

✿ Active Cognostics

⬛ Table Sort / Filter

📊 Univariate Filter
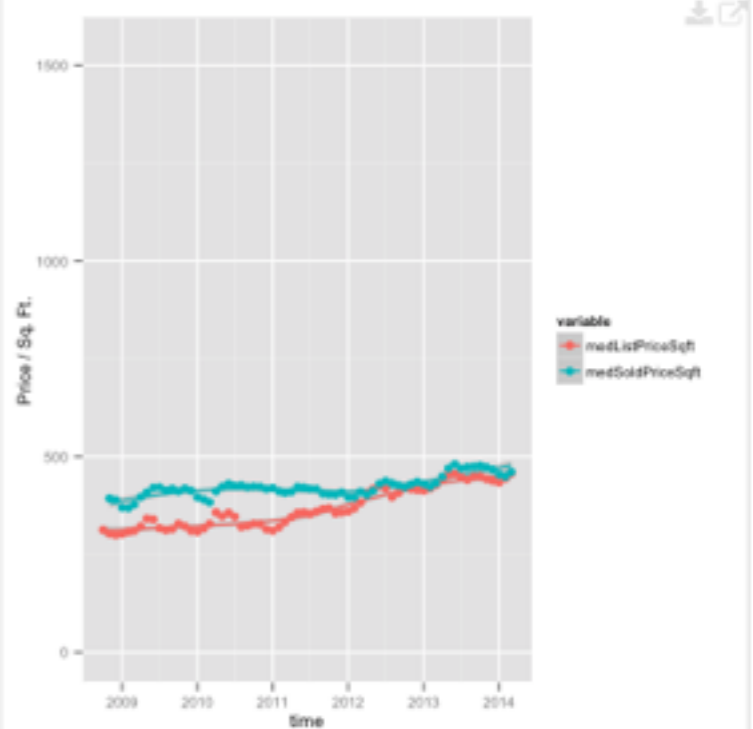
📈 Bivariate Filter

📉 Multivariate Filter

⠿ Sample

Cognostic	Show
panelKey	
panelKey: panel key	
condVar	
county: conditioning variable	✔
state: conditioning variable	✔
common	
slope: list price slope	✔
meanList: mean	✔
meanSold: mean	✔
nObs: number of non-NA list prices	✔
zillowHref: zillow link	✔

Cancel   Apply

State: ▼ slope ×   ▼ nObs ×   ⬇ slope ×

Future Work

- Split into 3 separate packages
    - trelliscopejs   - html / javascript
    - trelliscoper   - R server
    - trelliscopeServerjs - node.js server
- Functionality
    - Sample
    - Multi-panel display
    - Multivariate filtering
- Research
    - Automatic cognostics
    - Suggested ordering