

Introduction to Tableau Public

September 22, 2014

Angela Zoss (angela.zoss@duke.edu)

Data Visualization Coordinator

Data and Visualization Services, Perkins 226

<http://guides.library.duke.edu/tableau>

Today

Using Tableau Public 8.2 to create visualizations to explore a dataset:

- overview of the structure of the program
- terminology used
- sample data visualization project

To give feedback on this workshop:

<http://library.duke.edu/data/feedback>

What can Tableau make?

- Text tables
a grid representing variables by size and color
- Highlight tables
a grid representing variables by text and color
- **Maps (symbol, filled)**
- Pie charts
- Horizontal bars
- Stacked bars
- Side-by-side bars
- **Treemap**
a grid representing variables by size
- Circle views
- Side-by-side circles
- Lines/Area charts
- Lines/Area charts (discrete)
- Dual lines
- Dual combination
- Scatter plots
- **Histogram**
- **Box-and-whisker**
- **Gantt**
- **Bullet graphs**
- **Packed bubbles/Word cloud**



Tableau Public

- Free program
- Data limits
 - 1,000,000 rows of data
 - Excel, Access, text files
- Publishing limits
 - Can only save work to the **web**
(1GB storage space per account)
 - All data and views/dashboards will be **public**,
available for download

Tableau Desktop

- Allows you to save files privately
- Additional data connection options (databases, etc.)
- No data size limits

Free for:

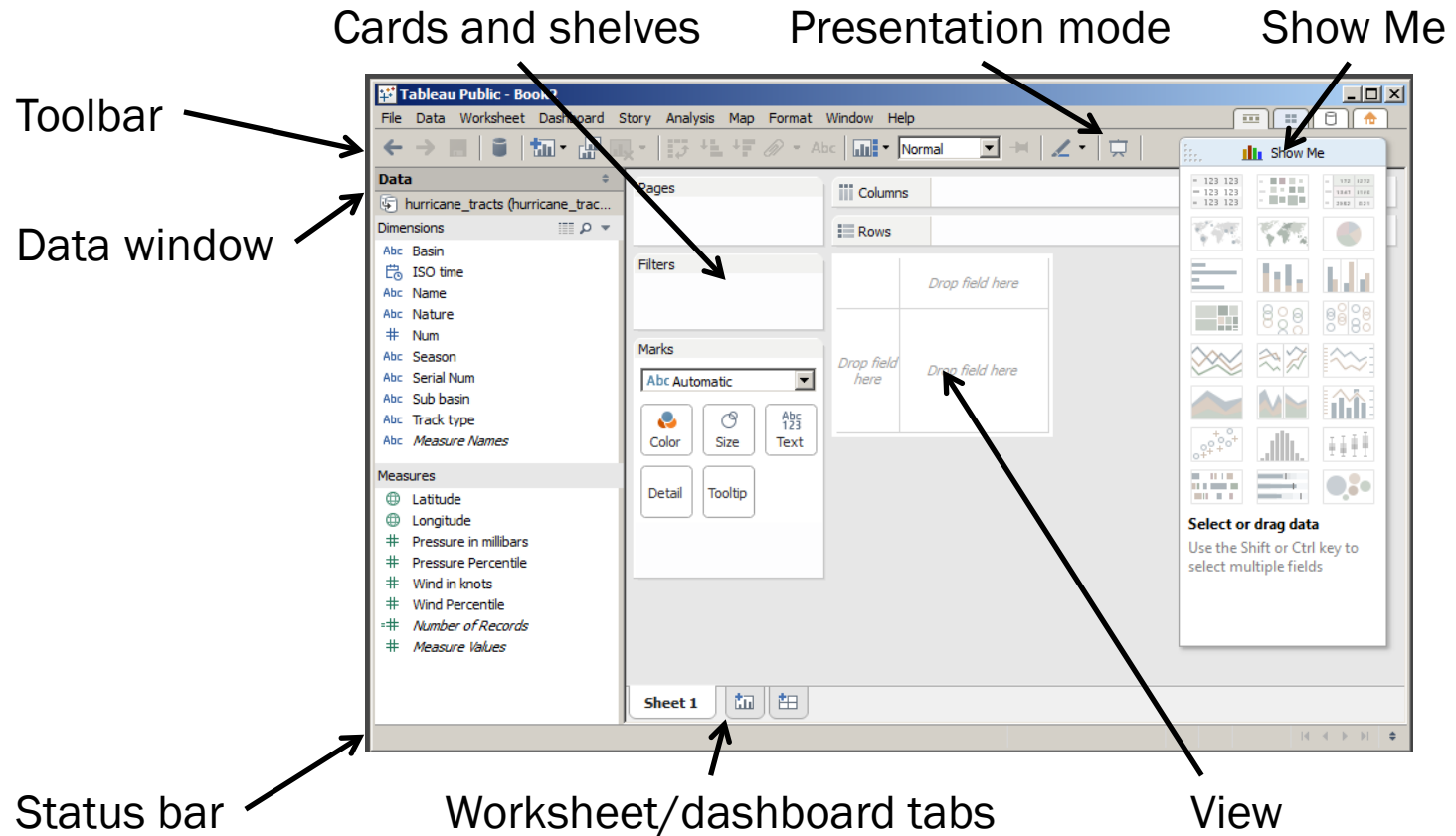
- 2 week trial
(<http://www.tableausoftware.com/products/trial>)
- students
(<http://www.tableausoftware.com/academic/students>)
- teachers using it in a class, semester license
(<http://www.tableausoftware.com/academic/teaching>)

Workbooks (Tableau .twbx files)

Contain one or more:

- Worksheets
 - Like a spreadsheet (a working space where data are organized/analyzed), but for creating visualizations (or “views”)
 - one (possibly complex) view per worksheet
- Dashboards
 - A presentation space where the views created in worksheets can be arranged and linked to produce a more complete visualization environment

The Tableau Workspace

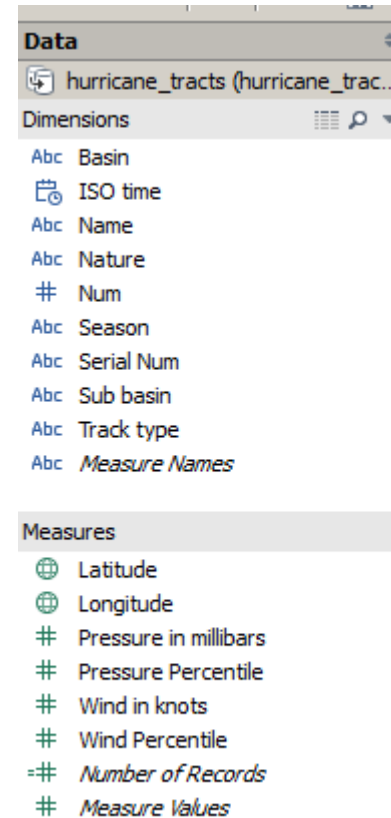


Online help:

<http://onlinehelp.tableausoftware.com/v8.2/public/online/en-us/help.htm>

Data Categories

- Dimensions
(text, date or boolean values)
- Measures
(numerical values)



<http://www.tableausoftware.com/public/knowledgebase/understanding-data-types-and-roles>

Tableau wants one column per variable

State	Sector	1990	1991	1992	1993
Alabama	Commercial	2.43	2.00	2.10	2.05
Alabama	Electric Power	50.28	54.18	56.93	62.82
Alabama	Industrial	25.15	25.36	28.36	26.18
Alabama	Residential	3.09	3.02	3.20	3.42
Alabama	Transportation	28.13	28.70	29.39	29.51
Alaska	Commercial	2.20	2.23	2.51	2.56
Alaska	Electric Power	2.61	2.46	2.29	2.32
Alaska	Industrial	15.83	17.44	18.87	18.51
Alaska	Residential	1.58	1.60	1.74	1.73
Alaska	Transportation	12.09	11.17	10.86	11.03
Arizona	Commercial	1.90	1.83	1.78	1.74
Arizona	Electric Power	32.52	32.76	35.37	36.58
Arizona	Industrial	3.86	3.92	3.93	3.94
Arizona	Residential	1.83	1.89	1.75	1.71
Arizona	Transportation	22.83	23.36	23.74	25.05
Arkansas	Commercial	1.62	1.63	1.59	1.81
Arkansas	Electric Power	21.60	21.74	22.11	19.80
Arkansas	Industrial	9.35	8.07	9.56	9.66
Arkansas	Residential	2.52	2.58	2.45	2.84
Arkansas	Transportation	16.16	16.30	16.22	17.02
California	Commercial	18.96	19.11	17.42	15.51
California	Electric Power	40.18	37.95	45.53	41.93
California	Industrial	71.62	73.00	73.29	70.53

State	Sector	Year	Emissions
Alabama	Commercial	1990	2.42906
Alabama	Commercial	1991	1.999039
Alabama	Commercial	1992	2.10271
Alabama	Commercial	1993	2.05046
Alabama	Commercial	1994	2.054954
Alabama	Commercial	1995	1.962636
Alabama	Commercial	1996	2.16117
Alabama	Commercial	1997	2.417254
Alabama	Commercial	1998	1.903054
Alabama	Commercial	1999	2.187003
Alabama	Commercial	2000	2.249259
Alabama	Commercial	2001	2.163411
Alabama	Commercial	2002	1.99245
Alabama	Commercial	2003	2.089619
Alabama	Commercial	2004	2.15691
Alabama	Commercial	2005	1.845364
Alabama	Commercial	2006	2.221402
Alabama	Commercial	2007	1.987872
Alabama	Commercial	2008	2.005586
Alabama	Commercial	2009	1.911421
Alabama	Commercial	2010	2.106

Preparing and Loading Data

- Excel formatting

<http://kb.tableausoftware.com/articles/knowledgebase/preparing-excel-files-analysis>

- Excel plug-in for reshaping, filling

<http://kb.tableausoftware.com/articles/knowledgebase/addin-reshaping-data-excel>

- Multiple tables (joins)

http://onlinehelp.tableausoftware.com/current/pro/online/en-us/help.htm#joins_addtables_multiple.html

EXAMPLE DATA SET: HURRICANE TRACTS

http://bit.ly/Tableau_hurricanes

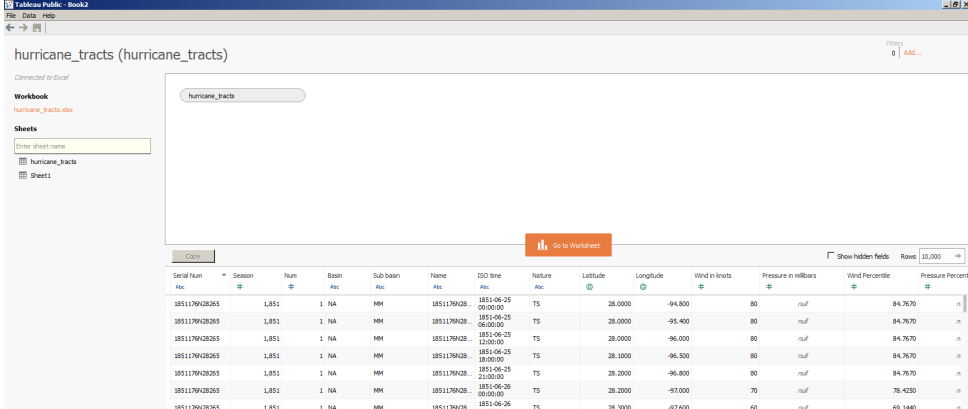
Source:

<http://www.ncdc.noaa.gov/ibtracs/index.php?name=wmo-data>

(downloaded July 2014)

Loading the data

- Open Tableau
- Click the “Open Data” button
- Select “Microsoft Excel”
- Select the
“hurricane_tracts.xlsx” file
- Drag “hurricane_tracts”
over *Drag sheets here*

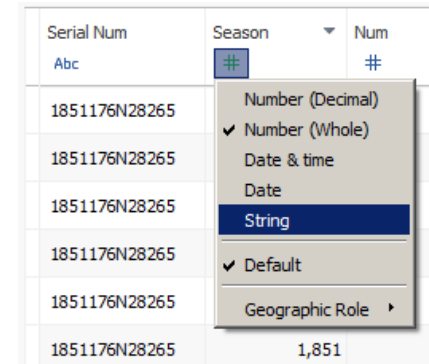


Serial Num	Season	Num	Basin	Sub basin	Name	ISO time	Nature	Latitude	Longitude	Wind in knots	Pressure in millibars	Wind Percentile	Pressure Percentile
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-25 00:00:00	TS	28.0000	-94.800	80	null	84.7670	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-25 06:00:00	TS	28.0000	-95.400	80	null	84.7670	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-25 12:00:00	TS	28.0000	-96.000	80	null	84.7670	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-25 18:00:00	TS	28.1000	-96.500	80	null	84.7670	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-25 21:00:00	TS	28.2000	-96.800	80	null	84.7670	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-26 00:00:00	TS	28.2000	-97.000	70	null	78.4250	in
1851179N28265	1,851	1	NA	NH	1851179N28	1851-06-26 06:00:00	TS	28.3000	-97.600	60	null	69.1440	in

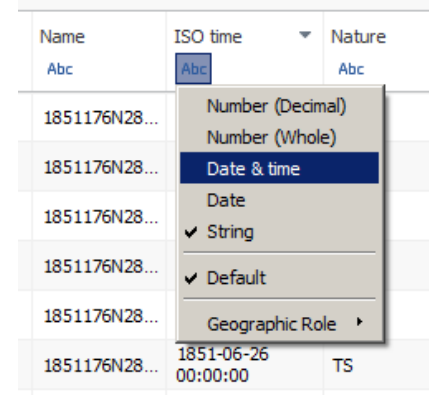
Fix variable type errors

- For “Season”, change variable type to String
- For “ISO time”, change variable type to Date & time

Serial Num	Season	Num
Abc	#	#
1851176N28265		
1851176N28265		
1851176N28265		
1851176N28265		
1851176N28265		
1851176N28265		
1851176N28265		1,851



Name	ISO time	Nature
Abc	Abc	Abc
1851176N28...		
1851176N28...		
1851176N28...		
1851176N28...		
1851176N28...		
1851176N28...		
1851176N28...	1851-06-26 00:00:00	TS



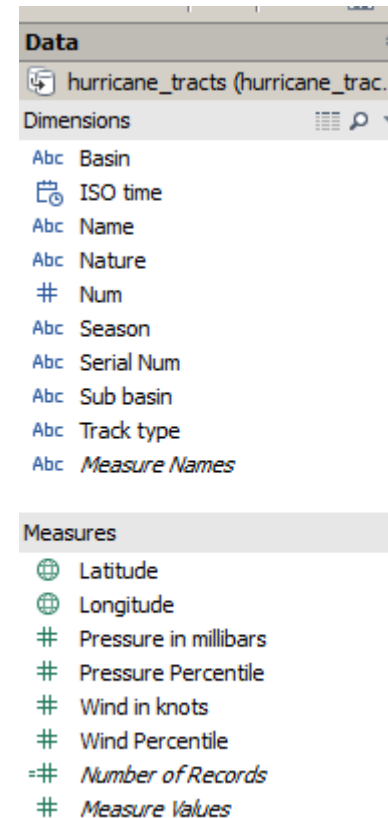
Tip: Years in Tableau

- Dates need a month and day
- If your data are annual, you can add a fake month and day that you later ignore

<http://community.tableausoftware.com/thread/113486>

Finalize import

- Click “Go to Worksheet”
- Move “Season (Count)” from Measures to Dimensions



VISUALIZING

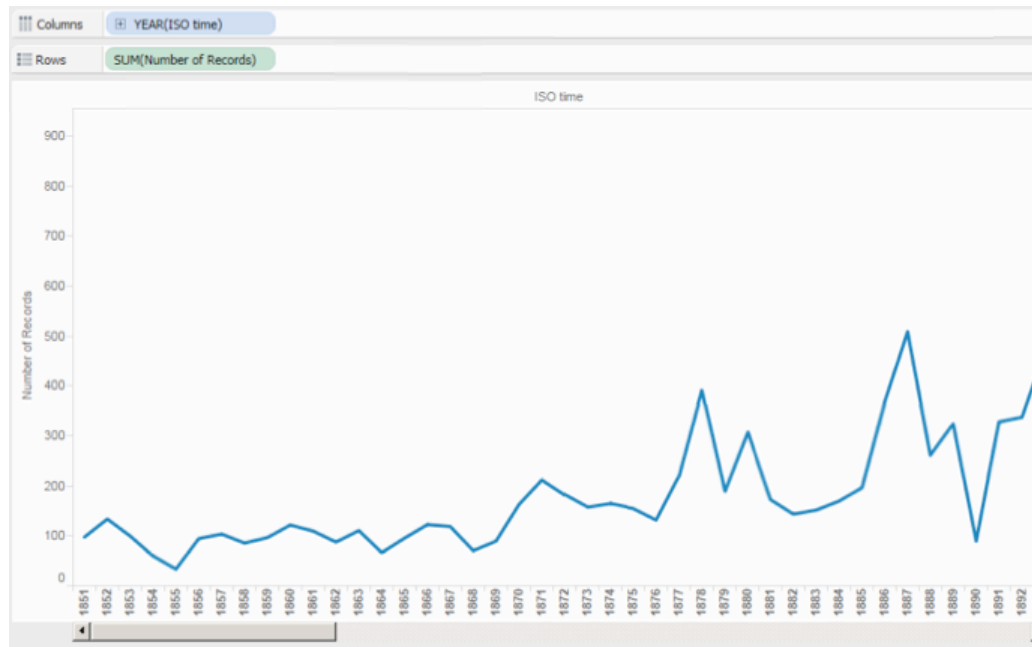
Building views

- Manually
 - drag and drop
- Using “Show Me!”
 - select desired dimensions and measures
 - click on “Show Me!”
 - mouse over suggestions for info



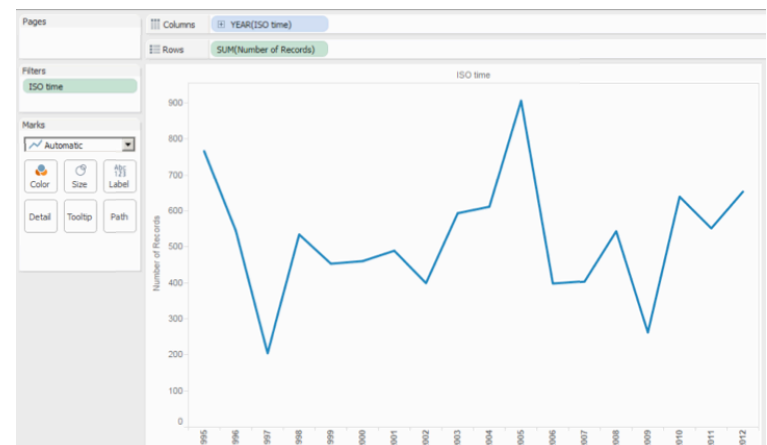
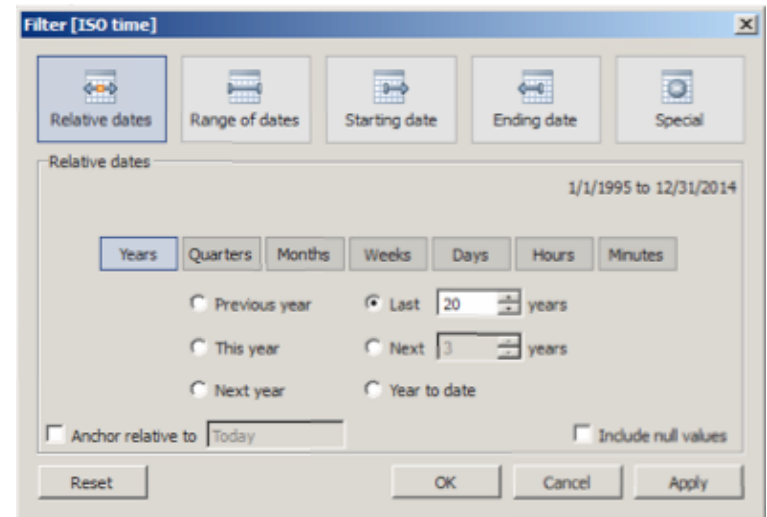
Create a line chart manually

- Drag “ISO time” to Columns or top of view area
- Drag “Number of Records” to Rows or left of view area



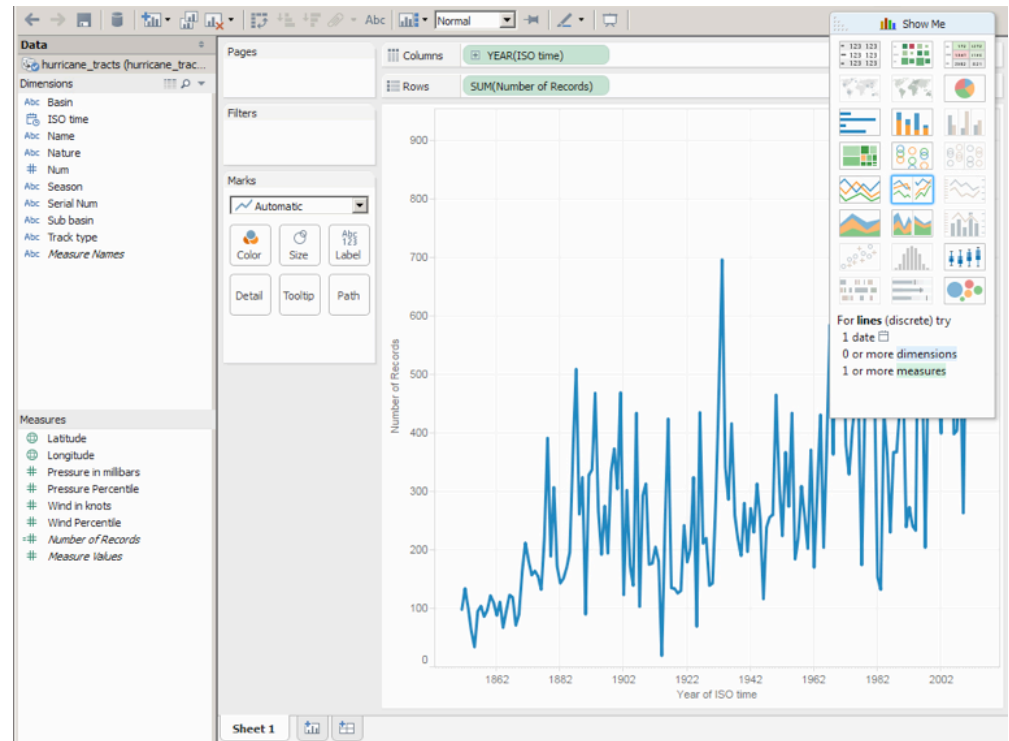
Limit to last 20 years

- Drag “ISO time” to Filters
- Click “Relative dates”
- Select “Years”
- Click the button for “Last”
- Adjust to 20 years



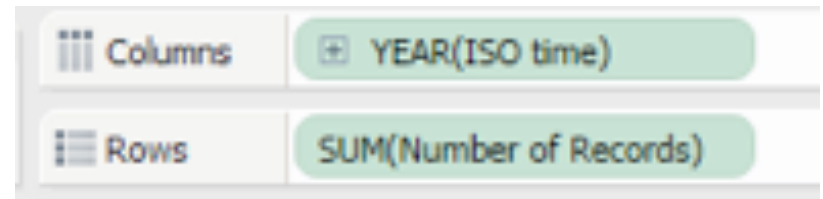
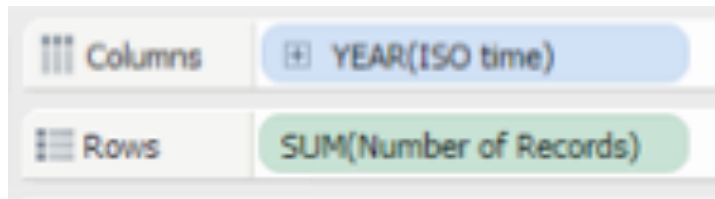
Create a line chart with Show Me

- Click “ISO time”
- Ctrl-click “Number of Records”
- Open Show Me
- Select line chart



Why are they so different?

NB: One has “Year(ISO time)” in **green**, while the other has it in **blue**.



Discrete versus continuous

- In the data window, “ISO time” is blue, which means **“discrete.”**

A discrete variable is like a category; values may have an order, but there isn't anything in between them.

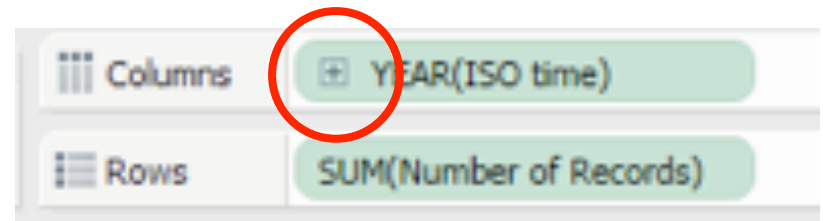
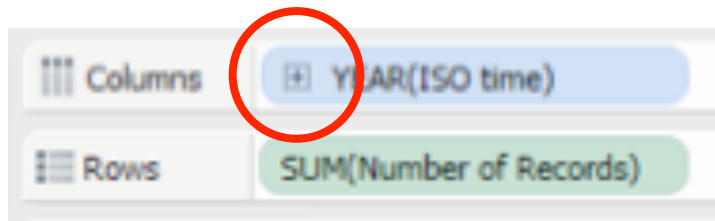
Discrete versus continuous

- In the data window, “ISO time” is blue, which means “discrete.”
- Line charts normally use a “continuous” date, though, so Show Me converted “ISO time” to green.

A continuous variable is like the full number range; you may only have a few values, but there are more possible numbers in between them.

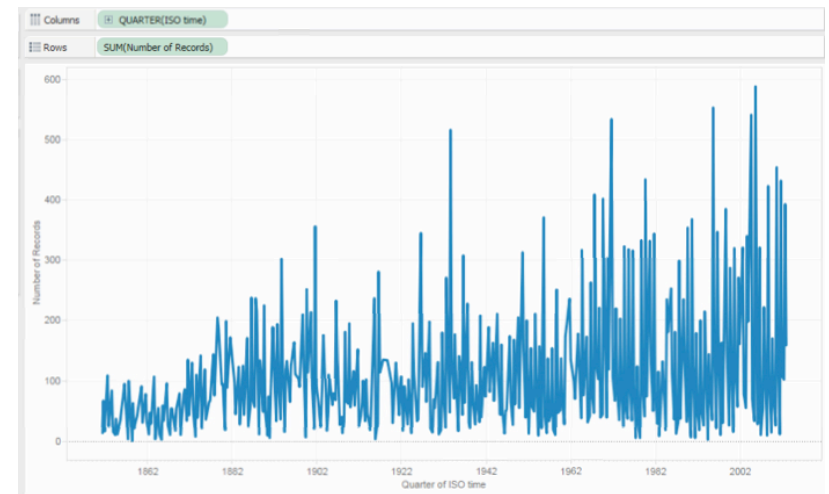
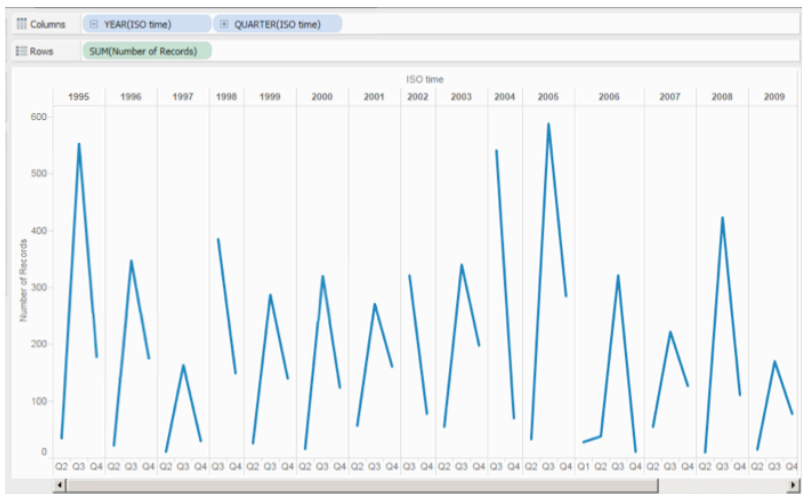
Discrete versus continuous

Try clicking the “+” sign inside the “Year(ISO time)” fields on both charts.



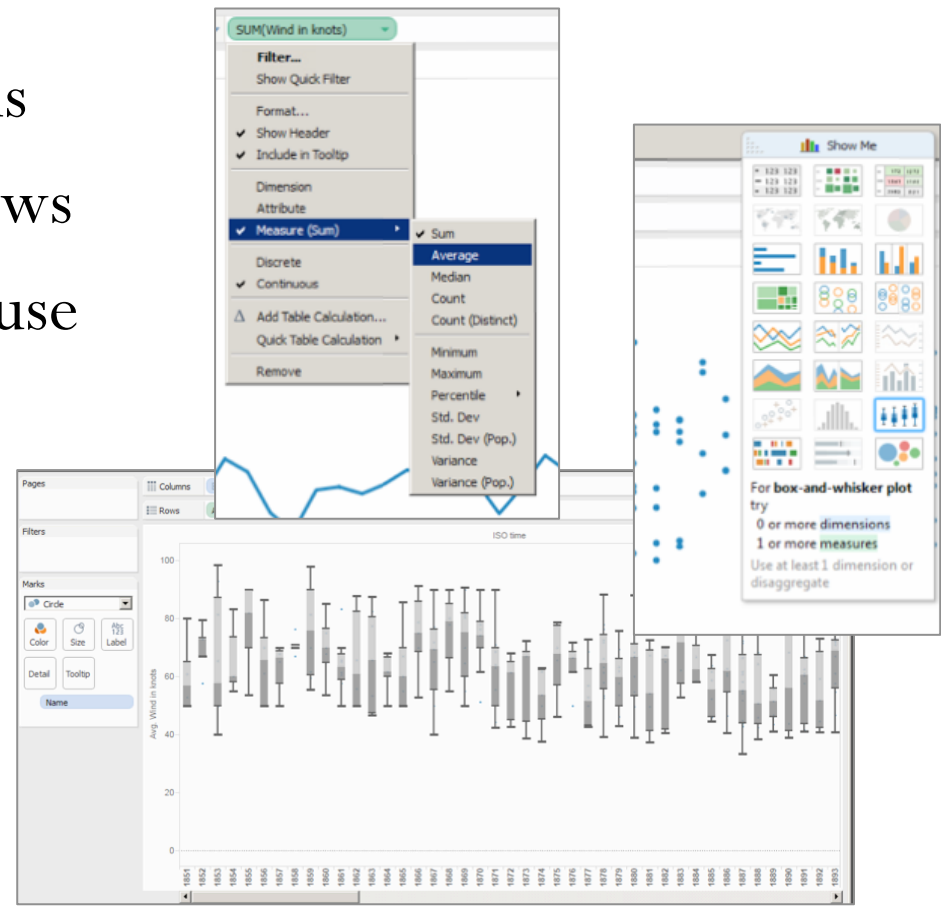
Discrete versus continuous

Both charts know that “Quarters” are a finer grain of detail than “Years”, but one treats them like categories, and the other just adds more points in between the years.



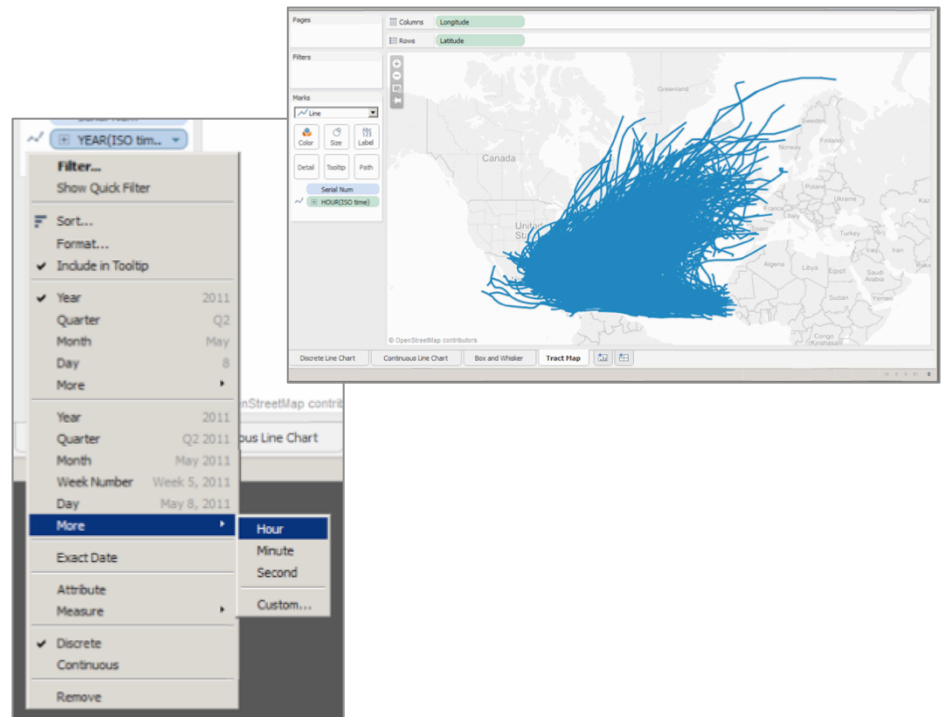
New in Tableau 8.2: Box-and-whisker plot

- Drag “ISO time” to Columns
- Drag “Wind in knots” to Rows
- Change “Wind in knots” to use Average instead of Sum
- Drag “Name” to Detail
- On Show Me, click on Box-and-whisker plot
- Limit to last 20 years



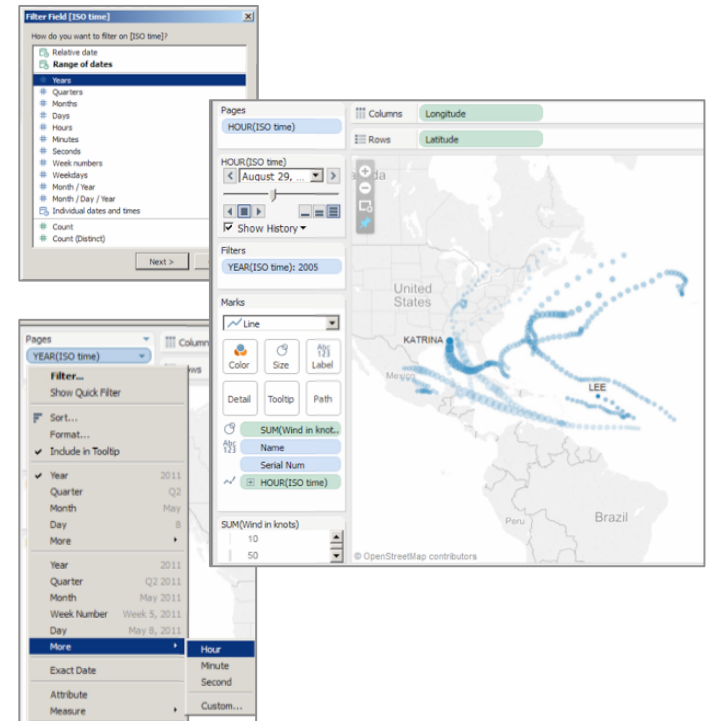
Mapping the tracts

- Double-click “Longitude” and “Longitude”
- Change both to Dimension
- Drag “Serial Num” to Detail
- Change mark type to Line
- Drag “ISO time” to Path
- Change “ISO time” to Continuous, Hour
- Optional: Drag “Wind in knots” to Size



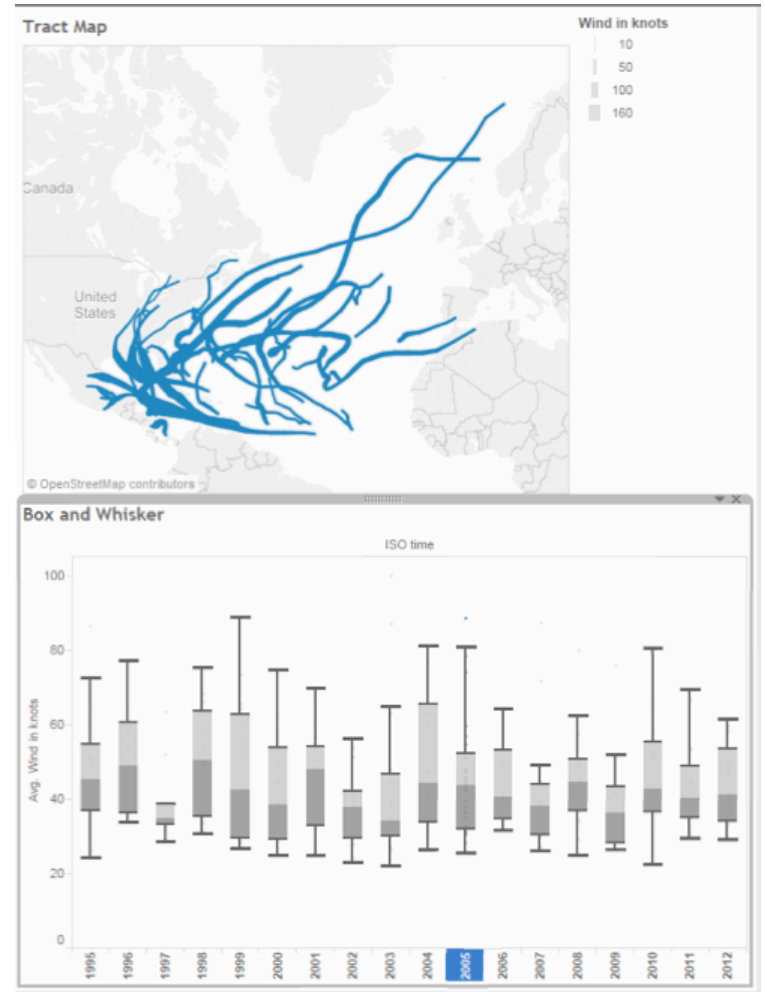
Animating the tracts

- Drag “ISO time” to Filters
- Click on Years (discrete) and select 2005 only
- Drag “ISO time” to Pages
- Click “Add all members”
- Change the Pages version of “ISO time” to Continuous, Hour
- Drag “Name” to Label
- Under Pages, click “Show History” and change settings to pick “All” marks
- Click the play button and speed up the animation



Create a dashboard

- Click the new dashboard button at the bottom
- Drag the unfiltered map to canvas
- Drag the box and whisker plot to the bottom of the canvas
- Click the arrow at the upper right of the box and whisker to select “Use as Filter”
- Click a year name or range to change the map



Saving

- Hide any sheets you don't want visible
- Click the “save” icon
- Log into your Tableau Public account (or create one)
- Choose a name for your workbook
- Decide if you want to show visible sheets as tabs
- Click Save
- Click “Open in Browser Window”
- Use Export button at bottom to get Image or PDF

GETTING HELP

Data and Visualization Services

- Data collections, LibGuides, etc.
<http://library.duke.edu/data/>
- Blog (tutorials, announcements, etc.)
<http://blogs.library.duke.edu/data/>
- Walk-in consultations
<http://library.duke.edu/data/about/schedule>
(or by appointment – askdata@duke.edu)
- Perkins 226 computing cluster
<http://library.duke.edu/data/about/lab>
(fast hardware, diverse software)
- Additional workshops
<http://library.duke.edu/data/news/>
(listserv – dvs-announce@duke.edu)

More on Data Visualization

What kinds of visualizations are out there?
How hard are they to reproduce?

See the Intro to Data Visualization LibGuide for general examples and tools/tutorials.

<http://guides.library.duke.edu/datavis/>

Additional Visualization LibGuides:

http://guides.library.duke.edu/vis_types

<http://guides.library.duke.edu/topten>

<http://guides.library.duke.edu/visualcomm>

QUESTIONS? SUGGESTIONS?

angela.zoss@duke.edu

<http://library.duke.edu/data/feedback>