Helping people see and understand data

Maureen Stone, Sr. Manager Tableau Research

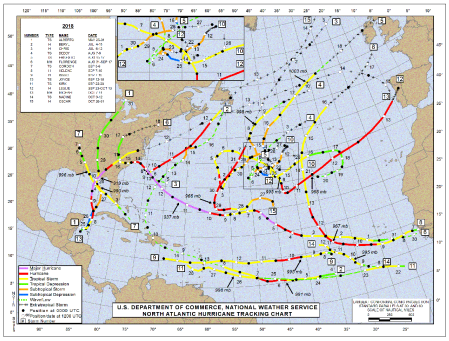
https://dfp.ubc.ca/news-and-events/events/helping-people-see-understand-data, includes a video

Roadmap

What does Tableau do?
Key technologies
Designing Tableau
Tableau Research

Tableau helps people see and understand data

Suppose you have data about hurricanes

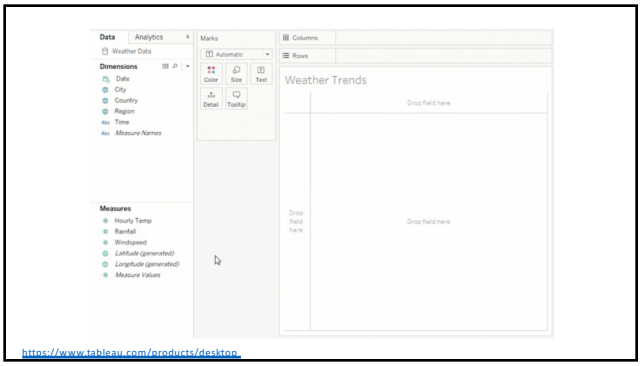


In the old days…

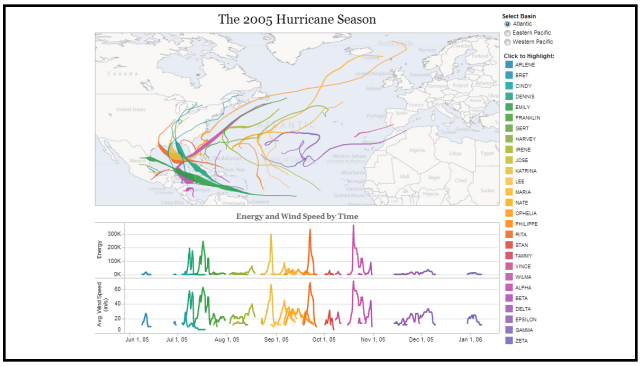
Write SQL to query your database
Use a graphing package to create a graph
Domain expert needs two other experts, at least!

With Tableau

Connect to your database
See your data schema
Use the Tableau GUI to explore and visualize your data



https://www.tableau.com/products/desktop



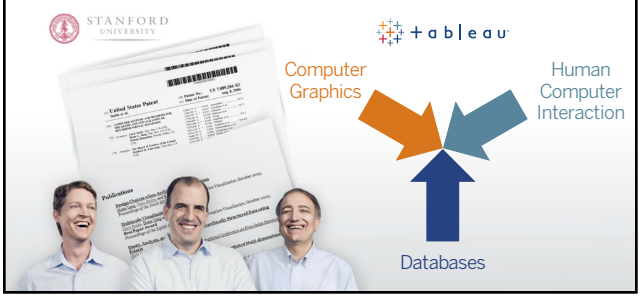
A few key points

Tableau is designed for enterprise data
- Large, aggregated
- Many different data sources

Tableau's target users are not data analysts
- Domain experts, people with the questions

Tableau design goals are challenging
- Easy to use, but supports deep analysis

Key technologies



The Polaris Formalism

"A formal language for describing table-based graphic presentations of multidimensional data that integrates analysis and visualization."

Stolte, Chris, Diane Tang, and Pat Hanrahan. "Polaris: A system for query, analysis, and visualization of multidimensional relational databases." *IEEE Transactions on Visualization and Computer Graphics* 8.1 (2002): 52-65.

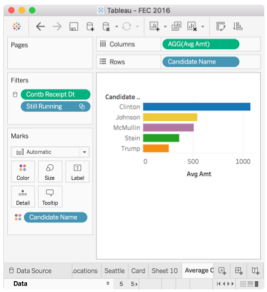
How this works

Pills on shelves
- Pills define what data
- Shelves describe layout intent

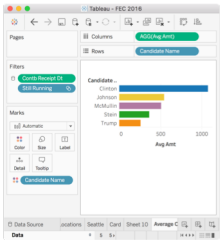
Generates a SQL query

Generates a visual specification

Visual spec rendered into a view

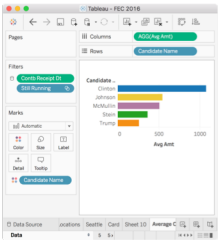


Generates a query

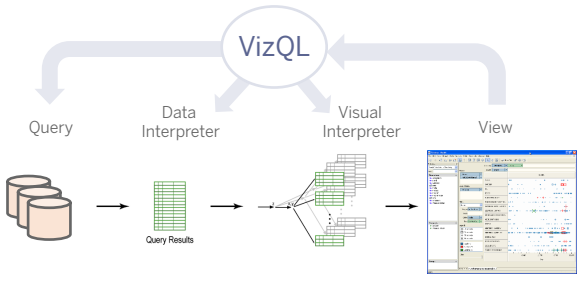


```
SELECT Candidate, AVG(Amount)
FROM FEC
WHERE Date > #2015-01-01#
  AND StillRunning is true
GROUP BY Candidate
```

Generates a visual spec



```
DISPLAY [AGG: Avg Amt]
ON ROWS,
[Candidate Name] ON
COLUMNS
[Candidate Name] ON
COLOR
AS BAR
FROM database
```



Stolte, Chris, Diane Tang, and Pat Hanrahan. "Polaris: a system for query, analysis, and visualization of multidimensional databases." *Communications of the ACM* 51.11 (2008): 75-84.

Allow people to easily and incrementally change the data they are looking at and how they are looking at it

Chris Stolte, Tableau co-founder

In addition…

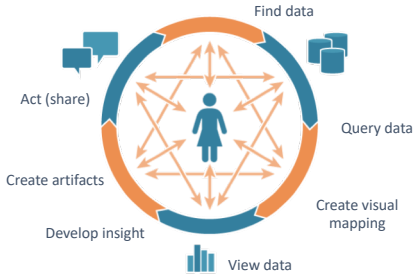
Data transformations
- Calcs, before and after the query

View transformations
- Layout, formatting
- Compose into multi-view dashboards

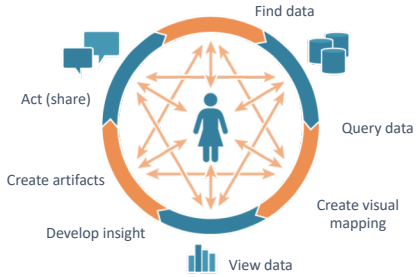
"Switzerland" of data

Designing Tableau

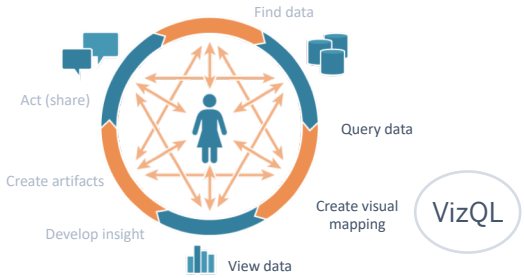
Cycle of analysis

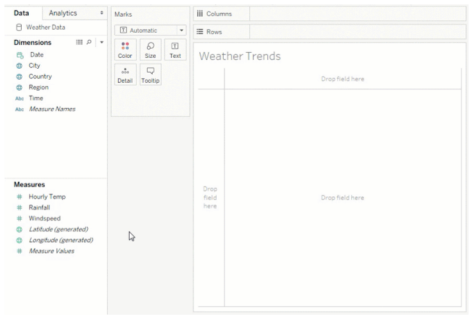


Task analysis: see and understand data



Analytic flow





https://www.tableau.com/products/desktop

Tableau UX design

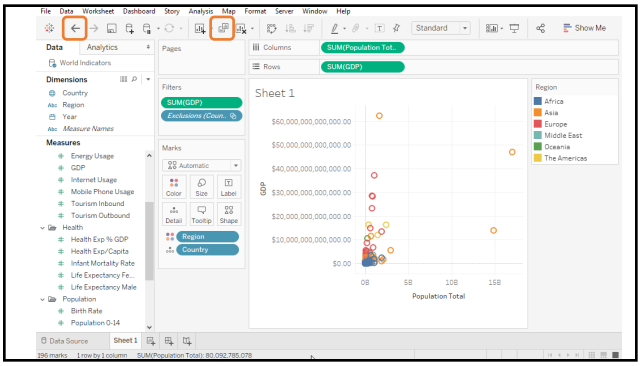
Incremental, Expressive, Unified, Direct, Effective

Important for the flow

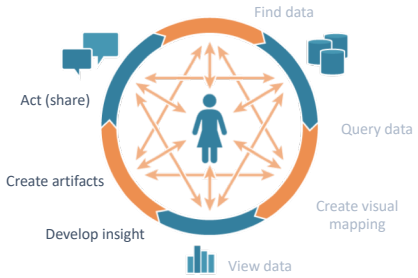
Show Me automatic presentation
- Automatic marks
- Rule-based recommendations
- Formal specifications

Mackinlay, Jock, Pat Hanrahan, and Chris Stolte. "Show me: Automatic presentation for visual analysis." *IEEE transactions on visualization and computer graphics* 13.6 (2007): 1137-1144.

Demo Show Me



Help people share their insights

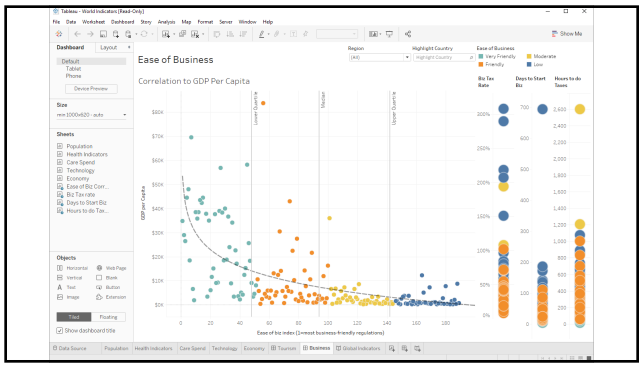


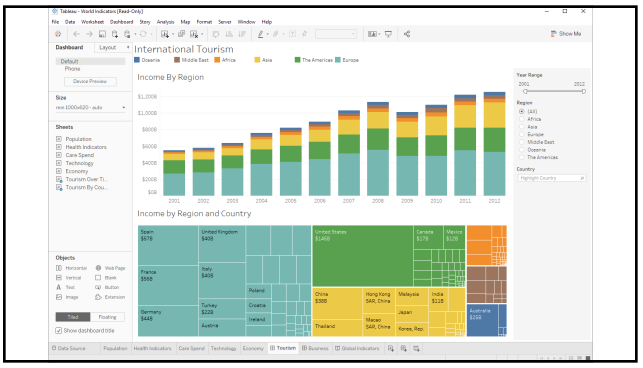
Key points

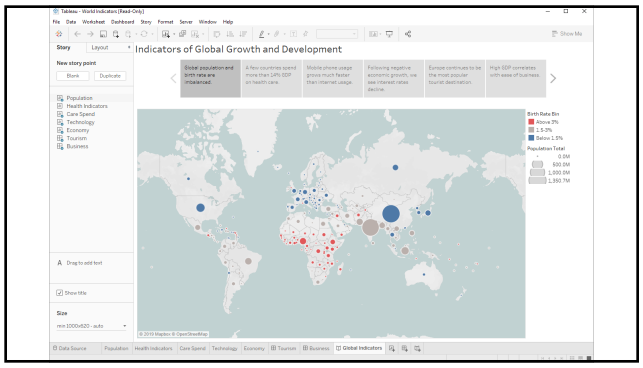
Tableau users create workbooks
- Views, dashboards, story points
- Data sources, embedded or separate

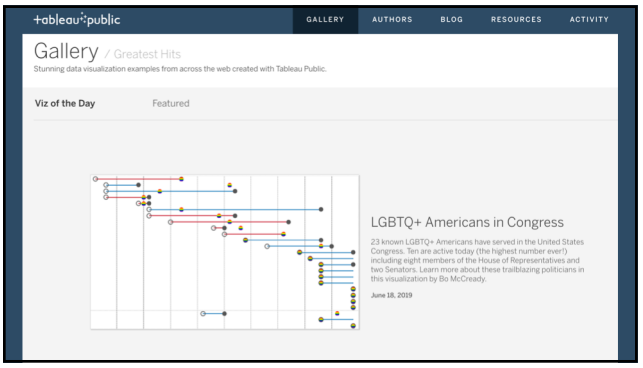
For other people
- Authors create and "publish" analytic artifacts
- "Consumers" work from these artifacts
- Reports, interactive dashboards, analytic applications

Good visual design is very important



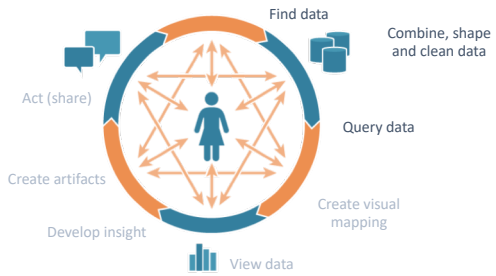




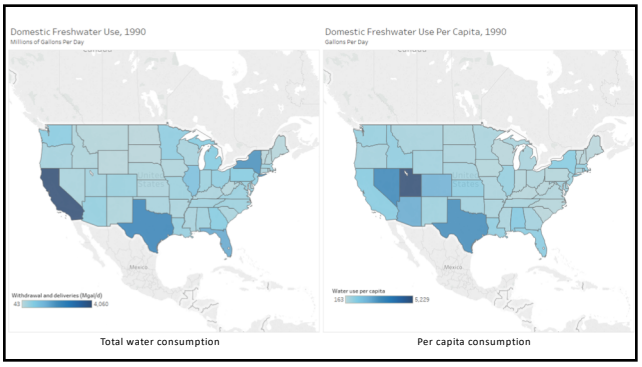


Fast, easy, beautiful

Help people get the right data



Need the right data to get the right answers



Total water consumption Per capita consumption

Tableau Prep (2018)

Data is rarely a simple table
- Multiple data sources
- Data shaping (joins, unions, pivots)
- Different shapes answer different questions

Data often needs “cleaning”
- Errors, wrong types, missing values

Tableau data is usually not static

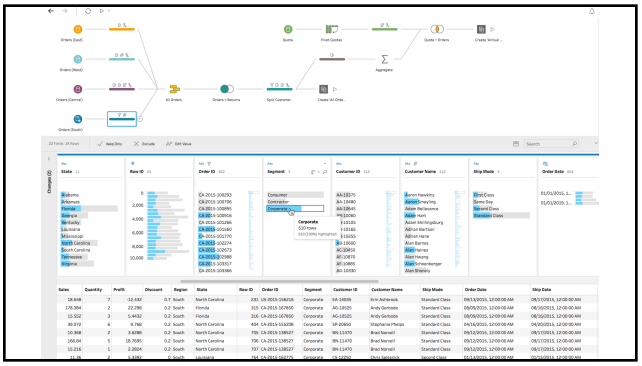


Tableau Prep creates data flows

Tableau Research

research.tableau.com

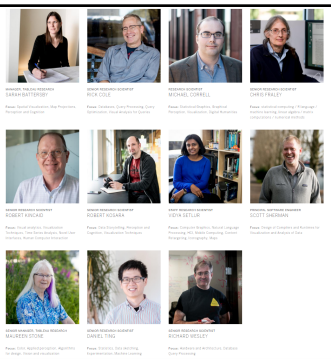
Tableau Research

History
- Started 2012, by Jock Mackinlay
  - Maureen Stone, Anushka Anand, Justin Talbot, Robert Kosara, Vidya Setlur
- 2017—part of RX (~11 people + Maureen as manager)
- 2018—part of OCTO (ditto)

Why?
- Tableau innovation based on academic research (Stolte's thesis)
- Continue this by creating an industrial research lab for Tableau

Visualization
Robert Kosara
Michael Correll
Scott Sherman
Robert Kincaid
Matthew Brehmer*
Anamaria Crisan*

Specialists
NLP: Vidya Setlur
Maps: Sarah Battersby
Color: Maureen Stone



Data Science & ML
Chris Fraley
Daniel Ting
(Ana Crisan)

Data systems
Rick Cole
Richard Wesley
(Daniel Ting)

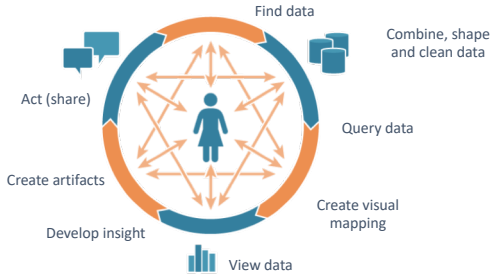
Interns
Michael Oppermann
Andrew McNutt
Moritz Sichert

What do we do?

We offer to Tableau
- Academic and prototyping research skills
- Domain-specific expertise, both technical and strategic
- Participation in the academic research community

That is…
- Read things, write things, build things
- Consult internally—both solving problems and limiting risk
- Make Tableau visible and influential; build our own skills/careers

Goal: Make all this easier, more effective



Some contributions…

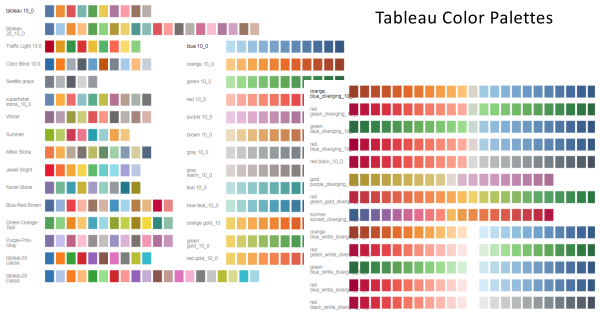
Color for data

Maureen Stone

Designed, and redesigned all of Tableau's data colors

Principles
- Functional yet beautiful
- Palettes and mappings, not color pickers
- Started in 2004, redesigned in 2010

Drove a research agenda in color for visualization



Color research

Color names

Setlur, Vidya, and Maureen C. Stone. "A linguistic approach to categorical color assignment for data visualization." *IEEE transactions on visualization and computer graphics* 22.1 (2015): 698-707.

Color and size

Stone, Maureen, Danielle Albers Szafir, and Vidya Setlur. "An engineering model for color difference as a function of size." *Color and Imaging Conference*. Vol. 2014. No. 2014. Society for Imaging Science and Technology, 2014.

Color affect

Bartram, Lyn, Abhisekh Patra, and Maureen Stone. "Affective color in visualization." *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017.

Storytelling

Robert Kosara

Tableau Story points feature

Tapestry Conference

Blogs, talks, podcasts, research papers

Kosara, Robert, and Jock Mackinlay. "Storytelling: The next step for visualization." *Computer* 46.5 (2013): 44-50.

Kosara, Robert. "Presentation-oriented visualization techniques." *IEEE computer graphics and applications* 36.1 (2016): 80-85.

Haroz, Steve, Robert Kosara, and Steve Franconeri. "ISOTYPE Visualization: Memory, Performance, and Engagement." (2018).

Query pipeline improvements

Analytic Query Language (AQL) — Justin Talbot

A high-level, strongly typed functional programming language that expresses all the computation required to produce the underlying data for rendering analytic views in Tableau.

Query-graph visualizer — Rick Cole

See and understand Tableau's query ecosystem (GitHub)

ML + Query optimization

Liqi Xu, Richard L. Cole, Daniel Ting. Learning to Optimize Federated Queries. To appear in aiDM'19, July 5, 2019, Amsterdam, Netherlands

Tableau "Ask Data"

Vidya Setlur & Melanie Tory

Started as a research project (Eviza, 2015)

- Create a research + dev team
- Acquire Cleargraph, build a bigger team

Ask Data feature released 2018

Ongoing NLP/NLI research

Setlur, Vidya, et al. "Eviza: A natural language interface for visual analysis." *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 2016.

Setlur, Vidya, and Melanie Tory. "Exploring Synergies between Visual Analytical Flow and Language Pragmatics." *2017 AAAI Spring Symposium Series*. 2017.

Setlur, Vidya, Melanie Tory, and Alex Djalali. "Inferencing Underspecified Natural Language Utterances in Visual Analysis." *Proceedings of the 24th International Conference on Intelligent User Interfaces*. ACM, 2019.

Vidya Setlur, Melanie Tory
Marti Hearst (visiting scientist)

What's coming next?

Get the right data

New data for new questions
- Text "wrangling" for semi-structured text
- Data sequences, events, intervals
- Relationship data (aka graphs)

Performance
- Query optimization
- Data sketching
- UDFs (with the Hyper team)

Enhanced analytic flow

Analytic "conversations"
- NLP, NLI + Visualization
- Understanding intent, pragmatics in this context

More automation
- Enhanced Show Me
- Suggestions, recommendations
- Data semantics

Increase understanding

Encourage skepticism
- Black hat visualization
- Visual summaries
- "Spell checking" for visualization

Integrate computational analytics
- Data science models (stats, ML)
- Human in the loop, not black box
- Answer and explain

Improve communication

Dashboards
- Who uses them and how?
- "Second cycle of analysis"

Presentation
- More visually expressive
- Easier to create and style
- Suitable for Tableau data

Tableau Vancouver

Tableau development office, downtown Vancouver

Head: Jesse Calderon

Org: Augmented Analytics
- Adding "AI" to Tableau's products
- Recommendations, Ask Data, Explain Data

Tableau Public

Two DFP projects already (Tamara)

300-545 Robson St., Vancouver, B.C.

Eric Brochu & Mya Warren

In summary

People need to understand their data
But understanding data is hard
Let's build tools to help them

We help people see and understand data