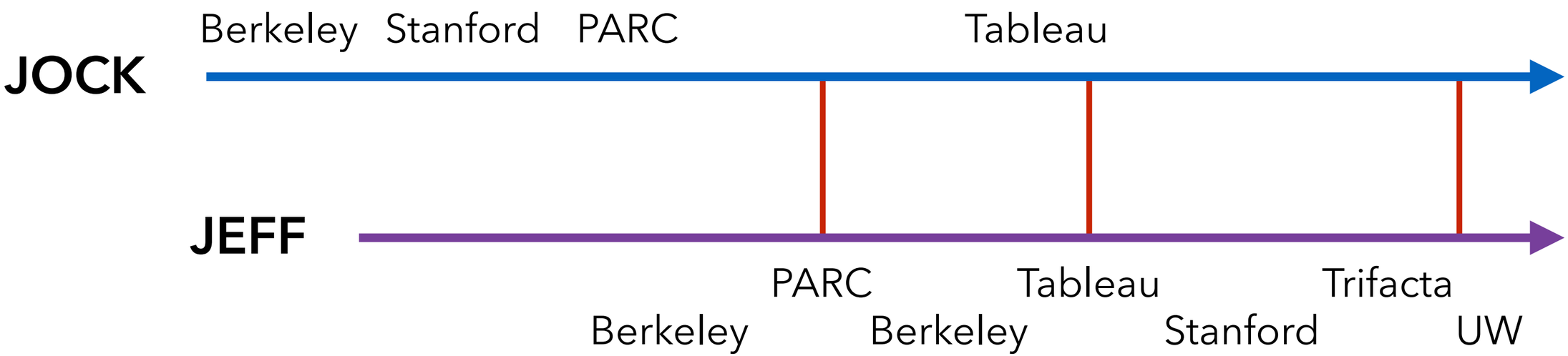
From Profiling to Analysis

DESIGNING VISUALIZATION TOOLS FOR PURPOSE

Jeffrey Heer
Trifacta / U. Washington

Jock Mackinlay
Tableau Software

By Way of Introduction…



A combined ~50 years of experience designing visualization tools!

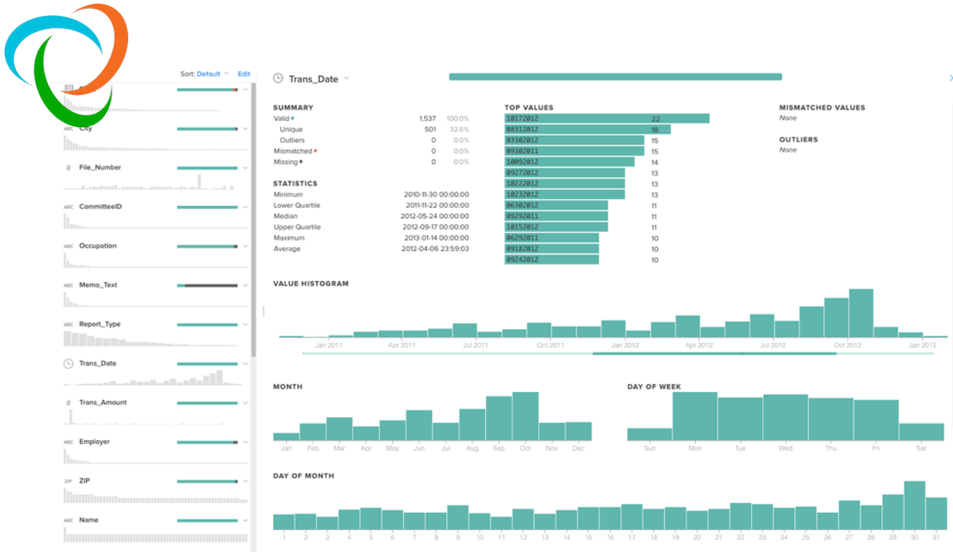
Designing for Purpose

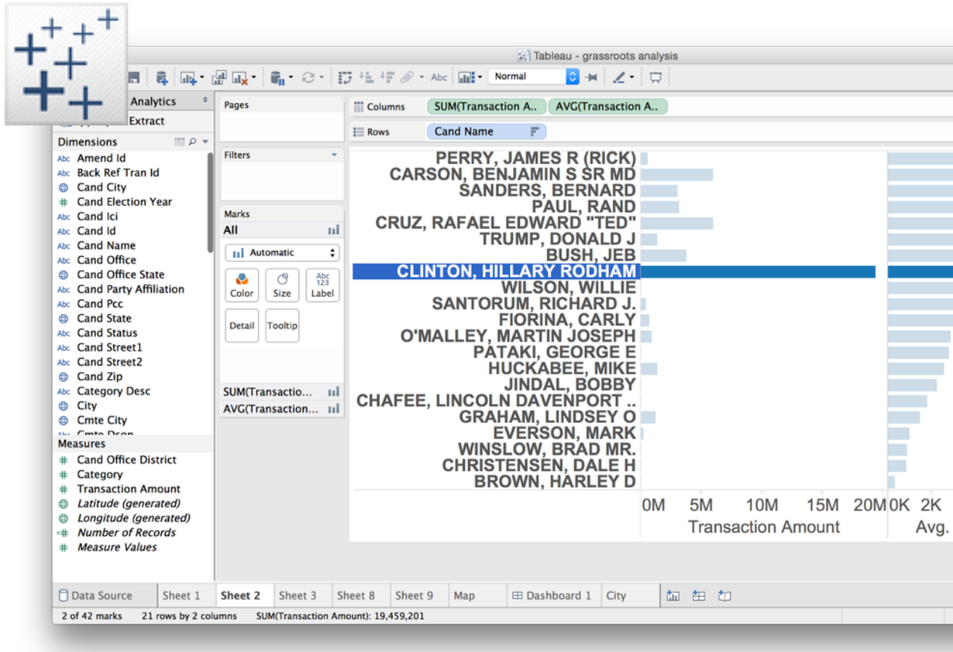
Visual analysis is central to both Trifacta & Tableau.

Each tool focuses on different complementary tasks, sometimes leading to divergent decisions for visualization and interaction design.

Here, we will examine how each tool uses visualization to help people make sense of data.

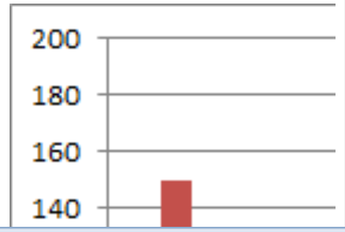
What lessons we can draw for the design of future visual analysis systems?

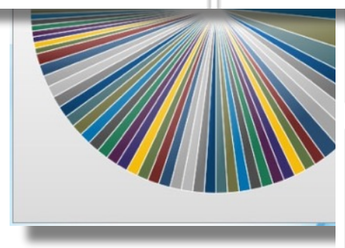


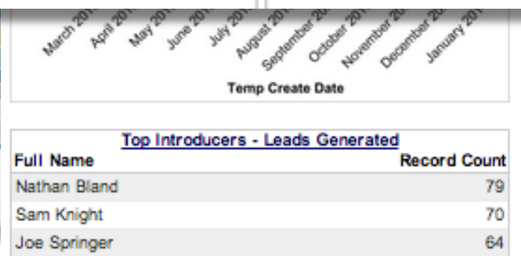


Leads

Leads Created MTD





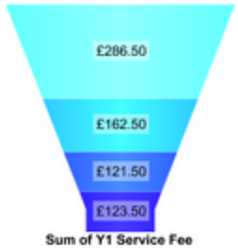


Top Introducers - Leads Generated

Full Name	Record Count
Nathan Bland	79
Sam Knight	70
Joe Springer	64
Frazer Smith	51
Andrew Davis	41
Adam Wallace	39

Deal Pipeline

Company Pipeline

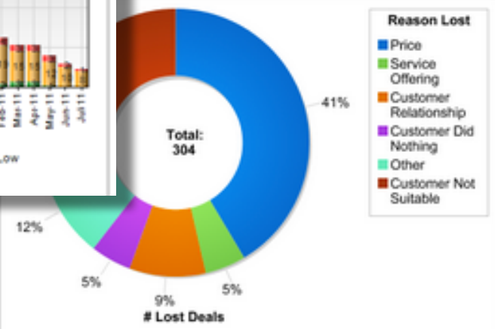


Sum of Y1 Service Fee (Thousands)

Stage
2 - Due Diligence
4 - Negotiation/Review

Pipeline by Rep

Lost Deals by Reason

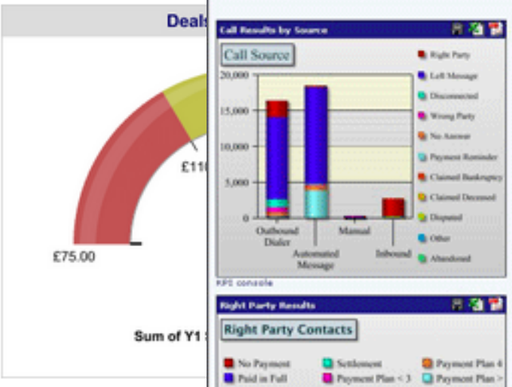


Current FY

Opportunity Conversion %

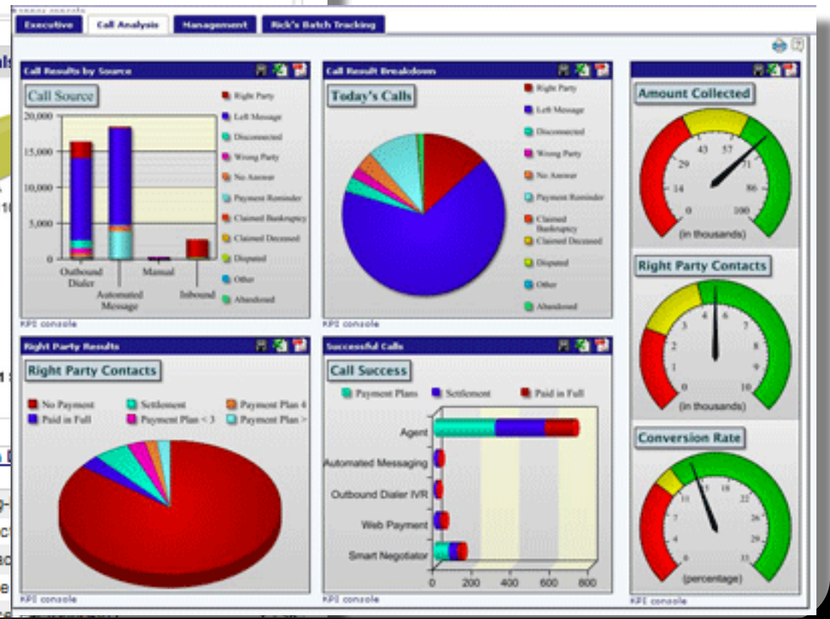
Opportunity Owner	Conversion Ratio
Caroline Swift	25
David Collins	22
Paul Jones	22

Deals Paid Out

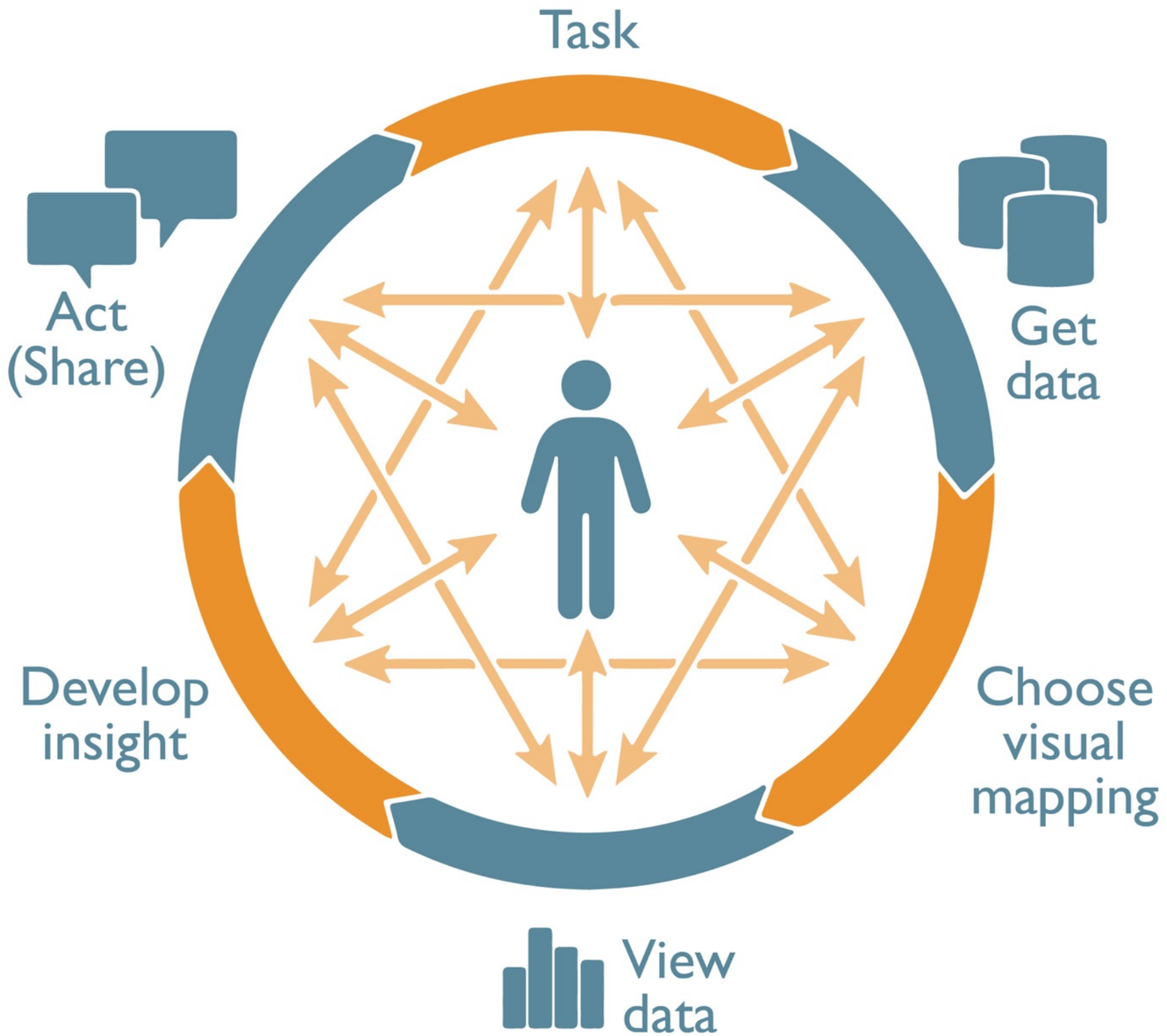


Won

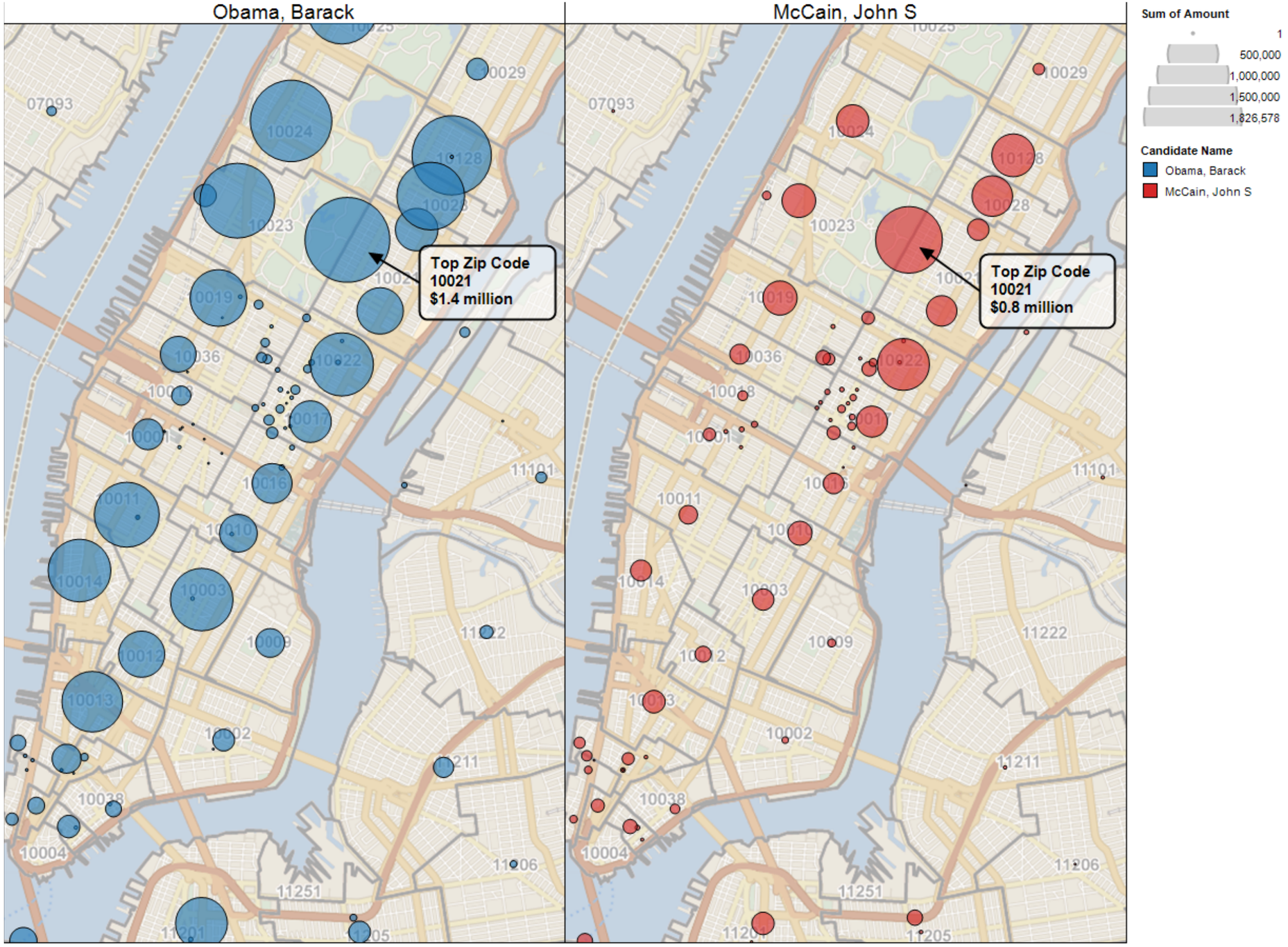
Opportunity Name
Tranter-Invoice Factoring-
Hedges Grow-Invoice Fac
Frink Security-Invoice Fac
Spencers Printing-Invoice
Sharp Engineering-Invoice



Task
Get data
Choose visual mapping
View data
Develop insight
Act (Share)


Task
Act
(Share)
Get
data
Develop
insight
Choose
visual
mapping
View
data

2008 Presidential Election: Where the Donors Are



Campaign Finance Disclosure Portal
Meetings and Hearings
Enforcement Matters
Help with Reporting and Compliance
Law, Regulations and Procedures
Commission Calendar

Detailed Files About Candidates, Parties and Other Committees

Files by Election Cycle

- 2015 - 2016
- 2013 - 2014
- 2011 - 2012
- Archive Files (1980 - 2010)

Tutorial: Working with Downloadable FEC Data Files Using MS Access.

These files contain committee, candidate and campaign finance data for the current election cycle and for election cycles through 1980. The files for the current election cycle plus the two most recent election cycles are regularly updated. The Committees, Candidates and Linkages files are updated daily. The Itemized Records, Contributions to Candidates, Individual Contributions and Operating Expenditures files are updated weekly on Sunday. When these files are updated they include data entered into the FEC database through that date. Please note for financial data complete entry from each reporting period takes about 30 days. It is important to be careful when making conclusions from these data because some information may not have completed the entry process when these files where created.

Committees (Committee Master)

The committee master file contains one record for each committee registered with the Federal Election Commission. This includes federal political action committees and party committees, campaign committees for presidential, house and senate candidates, as well as groups or organizations who are spending money for or against candidates for federal office.

The file contains basic information about the committees. The ID number the Commission assigned to the committee is first, along with the name of the committee, the sponsor, where appropriate, the treasurer's name and the committee's address. The file also includes information about what type of committee is being described, along with the candidate's ID number if it is a campaign committee. A comma delimited header file is available on this file's data dictionary page.

Candidates (Candidate Master)

The candidate master file contains one record for each candidate who has either registered with the Federal Election Commission or appeared on a ballot list prepared by a state elections office.

The file contains basic information about the candidate, including name, party, whether the candidate is an incumbent, challenger, or involved in an open seat, address, state and district in which the candidate is running and the year of the election for which the candidate is registered. (Note that incumbent/challenger status is dynamic in the current election cycle and there may be delays in identifying districts that will involve open seats. The file also includes the ID number assigned to the candidate by the FEC which is used in tracking campaign finance information about the campaign, as well as the ID number of the candidate's principal campaign committee. A comma delimited header file is available on this file's data dictionary page.

Linkages (Candidate-Committee Linkage)

This file contains one record for each candidate to committee linakge.

2015 - 2016 Data Files

Name	Data File	Total Records	Updated	Format Description
Committee Master File	cm16.zip	12916	29-SEP-2015	CM Data Dictionary
Candidate Master File	cn16.zip	4956	29-SEP-2015	CN Data Dictionary
Candidate Committee Linkage File	ccl16.zip	4741	29-SEP-2015	CCL Data Dictionary
Any Transaction from One Committee to Another	oth16.zip	118931	27-SEP-2015	OTH Data Dictionary
Contributions to Candidates (and other expenditures) from Committees	pas216.zip	72487	27-SEP-2015	PAS2 Data Dictionary
Contributions by Individuals	indiv16.zip	584347	27-SEP-2015	INDIV Data Dictionary
Operating Expenditures	oppexp16.zip	271021	27-SEP-2015	OPPEXP Data Dictionary

2013 - 2014 Data Files

Name	Data File	Total Records	Updated	Format Description
Committee Master File	cm14.zip	14899	29-SEP-2015	CM Data Dictionary
Candidate Master File	cn14.zip	5538	29-SEP-2015	CN Data Dictionary
Candidate Committee Linkage File	ccl14.zip	6248	29-SEP-2015	CCL Data Dictionary
Any Transaction from One Committee to Another	oth14.zip	569066	27-SEP-2015	OTH Data Dictionary
Contributions to Candidates (and other expenditures) from Committees	pas214.zip	350857	27-SEP-2015	PAS2 Data Dictionary
Contributions by Individuals	indiv14.zip	2132431	27-SEP-2015	INDIV Data Dictionary
Operating Expenditures	oppexp14.zip	1347425	27-SEP-2015	OPPEXP Data Dictionary

2011 - 2012 Data Files

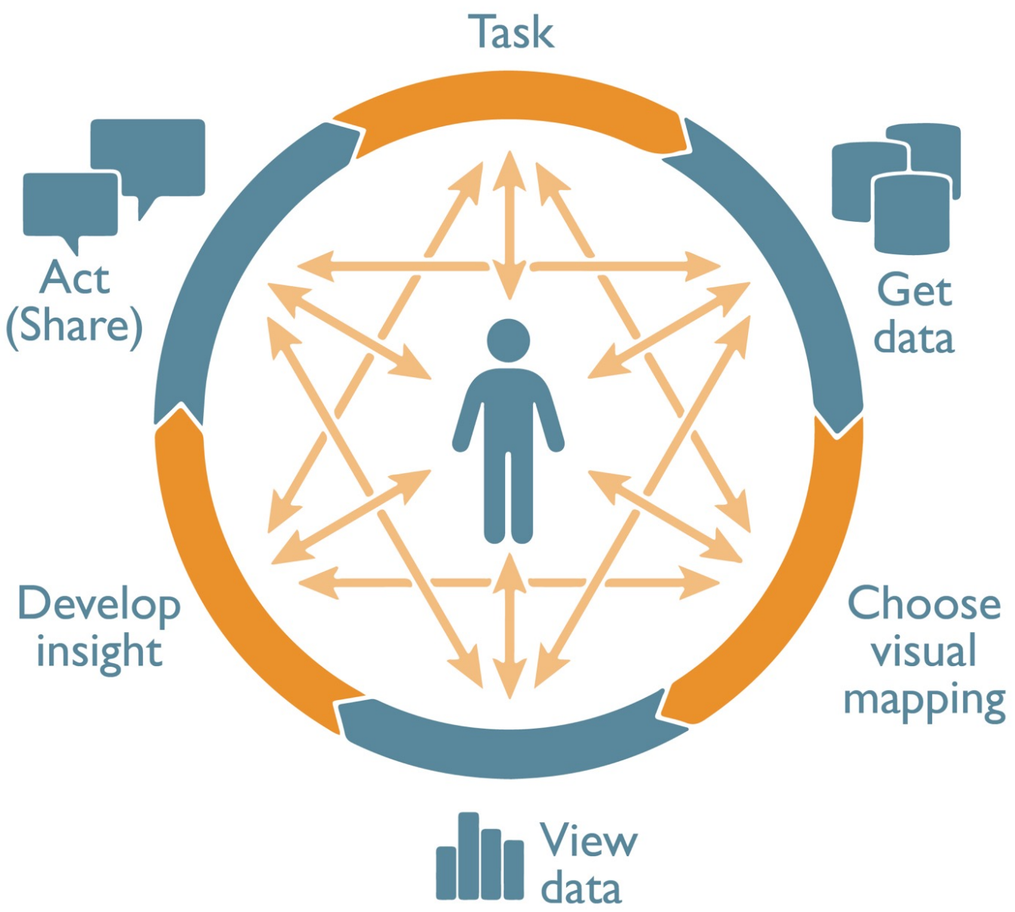
Name	Data File	Total Records	Updated	Format Description
Committee Master File	cm12.zip	14453	29-SEP-2015	CM Data Dictionary
Candidate Master File	cn12.zip	5628	29-SEP-2015	CN Data Dictionary
Candidate Committee Linkage File	ccl12.zip	6131	29-SEP-2015	CCL Data Dictionary
Any Transaction from One Committee to Another	oth12.zip	1012156	27-SEP-2015	OTH Data Dictionary
Contributions to Candidates (and other expenditures) from Committees	pas212.zip	396425	27-SEP-2015	PAS2 Data Dictionary

oppexp.txt | cm.txt | cn.txt

```
CMTE_ID|AMEND_IND|REPORT_YEAR|REPORT_TYPE|IMAGE_NUM|LINE_NUM|FORM_TYPE_CODE|SCHED_TYPE_CODE|NAME|CITY|STATE|ZIP_CODE|TRANSACTION_DATE|TRANSACTION_AMOUNT|TRANSACTION_PGI|PURPOSE|CATEGORY|CATEGORY_DESC|MEMO_CODE|MEMO_TEXT|ENTITY_TYPE|SUB_ID|FILE_NUM|TRAN_ID|BACK_REF_TRAN_ID
C00004606|N|2015|M4|15951124870|21B|F3X|SB|DIRECT MAIL SYSTEMS, INC.|CLEARWATER|FL|33762|03/13/2015|600||POSTAGE|003|Solicitation and Fundraising Expenses|||ORG|4041320151241802165|1002259|SB21B.20726||
C00004606|N|2015|M4|15951124870|21B|F3X|SB|PLAZA 21 LLC|TOPEKA|KS|66675|03/27/2015|982.8||RENT|001|Administrative/Salary/Overhead Expenses|||ORG|4041320151241802166|1002259|SB21B.20801||
C00378067|N|2015|Q1|15951124891|21B|F3X|SB|JPMORGAN CHASE BANK|NEW YORK|NY|10165|03/31/2015|233.96||||||ORG|4041320151241802167|1002260|SB21B.4776||
C00010470|N|2015|M4|15970338179|21B|F3X|SB|WELLS FARGO BANK NEBRASKA, N.A.|OMAHA|NE|68103|03/31/2015|269.37||BANK SERVICE FEE|001|Administrative/Salary/Overhead Expenses||BANK SERVICE FEE|ORG|4041320151241802266|1002270|22403718||
C00554774|N|2015|Q1|15951124928|21B|F3X|SB|PICKETT, MEGAN|RAYMOND|MS|39154|02/19/2015|400|G|CANVASSING / DOOR-TO-DOOR GOTV|001|Administrative/Salary/Overhead Expenses|||IND|4041320151241802267|1002271|SB21B.5013||
C00465211|N|2015|M4|15951124991|21B|F3X|SB|POLITICAL ACTION COMMITTEE SERVICES LLC|CLINTON|MD|20735|03/01/2015|750||PAC ADMINISTRATIVE EXPENSE, SUPPLIES, RENT, POSTAGE|001|Administrative/Salary/Overhead Expenses ||PAC ADMINISTRATIVE EXPENSE, SUPPLIES, RENT, POSTAGE|ORG|4041320151241802314|1002280|37970601||
C00374363|N|2015|M4|15951125105|21B|F3X|SB|SUN TRUST BANK|WASHINGTON|DC|20037|03/10/2015|24.95||MERCHANT FEE|001|Administrative/Salary/Overhead Expenses|||ORG|4041320151241802338|1002290|SB21B.19530||
C00374363|N|2015|M4|15951125105|21B|F3X|SB|SUN TRUST BANK|WASHINGTON|DC|20037|03/19/2015|343.72||ACCOUNT ANALYSIS FEE|001|Administrative/Salary/Overhead Expenses|||ORG|4041320151241802339|1002290|SB21B.19531||
C00382028|A|2015|M5|15971210806|21B|F3X|SB|T-MOBILE|ALBUQUERQUE|NM|87176|04/09/2015|88.78||TELEPHONE|||||ORG|4061720151245761887|1010651|3270||
C00008169|N|2015|Q1|15970340702|21B|F3X|SB|GRAVES AND ASSOCIATES|JEFFERSON CITY|MO|65109|03/05/2015|310||TAX PREPARATION|001|Administrative/Salary/Overhead Expenses|||ORG|4041420151241820124|1002618|SB21B.4161||
C00008169|N|2015|Q1|15970340702|21B|F3X|SB|INTERNAL REVENUE SERVICE||||02/24/2015|550||TAX PAYMENT|001|Administrative/Salary/Overhead Expenses|||ORG|4041420151241820125|1002618|SB21B.4159||
C00445593|N|2015|Q1|15951129812|21B|F3X|SB|WELLS FARGO|BOULDER|CO|80303|01/12/2015|14.75|G|BANK FEE|001|Administrative/Salary/Overhead Expenses ||BANK FEE|ORG|4041420151241820127|1002623|SB21B-EX123||
C00445593|N|2015|Q1|15951129812|21B|F3X|SB|WELLS FARGO|BOULDER|CO|80303|02/11/2015|20.29|G|BANK FEE|001|Administrative/Salary/Overhead Expenses ||BANK FEE|ORG|4041420151241820128|1002623|SB21B-EX124||
C00445593|N|2015|Q1|15951129812|21B|F3X|SB|WELLS FARGO|BOULDER|CO|80303|03/11/2015|20.18|G|BANK FEE|001|Administrative/Salary/Overhead Expenses ||BANK FEE|ORG|4041420151241820129|1002623|SB21B-EX125||
C00570226|N|2015|Q1|15970340711|21B|F3X|SB|BOWDITCH & DEWEY, LLP|BOSTON|MA|02110|03/30/2015|1750|G|LEGAL CONSULTING|||||ORG|4041420151241820148|1002630|SB21B.4407||
C00570226|N|2015|Q1|15970340711|21B|F3X|SB|SHAWMUT GROUP|CHESTNUT HILL|MA|02467|03/30/2015|1750|G|POLITICAL CONSULTING|||||ORG|4041420151241820149|1002630|SB21B.4404||
C00570226|N|2015|Q1|15970340711|21B|F3X|SB|SWIFTKURRENT|BEVERLY|MA|01915|03/30/2015|1227.43|G|SOCIAL MEDIA SERVICES|||||ORG|4041420151241820150|1002630|SB21B.4402||
C00570226|N|2015|Q1|15970340712|21B|F3X|SB|SWIFTKURRENT|BEVERLY|MA|01915|03/30/2015|1750|G|SOCIAL MEDIA SERVICES|||||ORG|4041420151241820151|1002630|SB21B.4405||
C00393769|N|2015|M4|15970340878|21B|F3X|SB|WELLS FARGO|PORTLAND|OR|97228|03/11/2015|124.95||CREDIT CARD PROCESSING FEES|||||ORG|4041420151241820190|1002653|SB21B.4123||
C00393769|N|2015|M4|15970340878|21B|F3X|SB|WELLS FARGO|PORTLAND|OR|97228|03/31/2015|5||BANK FEE|||||ORG|4041420151241820191|1002653|SB21B.4124||
C00043992|N|2015|M4|15951130193|21B|F3X|SB|MERCHANT SERVICES|KNOXVILLE|TN|37920|03/02/2015|89.94||CREDIT CARD PROCESSING FEES|||||ORG|4041420151241820198|1002657|SB21B.4114||
C00473918|N|2015|M4|15951130405|21B|F3X|SB|AUTHORIZE.NET|AMERICAN FORK|UT|84003|03/03/2015|80||CREDIT CARD SERVICE CHARGES|||||ORG|4041520151241876607|1002689|SB21B-653||
C00473918|N|2015|M4|15951130405|21B|F3X|SB|FIRST DATA MERCHANT SERVICES|HAGERSTOWN|MD|21741|03/03/2015|25||CREDIT CARD SERVICE CHARGES|||||ORG|4041520151241876608|1002689|SB21B-654||
C00473918|N|2015|M4|15951130405|21B|F3X|SB|AMERICA VOTES|WASHINGTON|DC|20036|03/06/2015|25000||PUBLICATION & DUES|||||ORG|4041520151241876609|1002689|SB21B-651||
C00473918|N|2015|M4|15951130406|21B|F3X|SB|HART RESEARCH ASSOCIATES|WASHINGTON|DC|20009|03/19/2015|100000||POLLING/SURVEYS|||||ORG|4041520151241876610|1002689|SB21B-652||
C00040253|N|2015|M4|15951130487|21B|F3X|SB|FIFTH THIRD BANK|ELMHURST|IL|60062|03/11/2015|85.76|O|SERVICE CHARGE|001|Administrative/Salary/Overhead Expenses|||ORG|4041520151241876611|1002691|B550588||
C00553271|N|2015|Q1|15951130719|21B|F3X|SB|CREATIVE CONSULTING|DEER PARK|NY|11729|01/02/2015|2540|P|SOCIAL MEDIA CONSULTING|||||ORG|4041420151241820330|1002703|SB21B.4226||
C00417295|N|2015|Q1|15951130847|21B|F3X|SB|247 CAMBRIDGE STREET TRUST|BELMONT|MA|02478|02/03/2015|350||RENT/UTILITIES FOR MASS PAC OCAMPAIGN OFFICE|001|Administrative/Salary/Overhead Expenses |||ORG|4041420151241820360|1002717|SB21B.5145||
```

TRIFACTA

Task: Data Wrangling



Automatic

Effective

Scalable

Actionable

Automatic

Sort: Default Edit

- State
- City
- File_Number
- CommitteeID
- Occupation
- Memo_Text
- Report_Type
- Trans_Date
- Trans_Amount
- Employer
- ZIP
- Name

Trans_Date

SUMMARY

Valid	1,537	100.0%
Unique	501	32.6%
Outliers	0	0.0%
Mismatched	0	0.0%
Missing	0	0.0%

STATISTICS

Minimum	2010-11-30 00:00:00
Lower Quartile	2011-11-22 00:00:00
Median	2012-05-24 00:00:00
Upper Quartile	2012-09-17 00:00:00
Maximum	2013-01-14 00:00:00
Average	2012-04-06 23:59:03

TOP VALUES

10172012	22
08312012	18
03302012	15
09302011	15
10092012	14
09272012	13
10222012	13
10232012	13
06302012	11
09292011	11
10152012	11
06292011	10
09182012	10
09242012	10

MISMATCHED VALUES

None

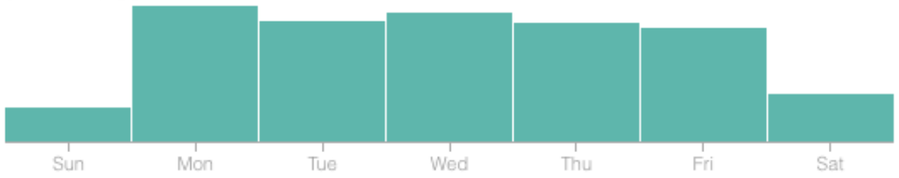
OUTLIERS

None

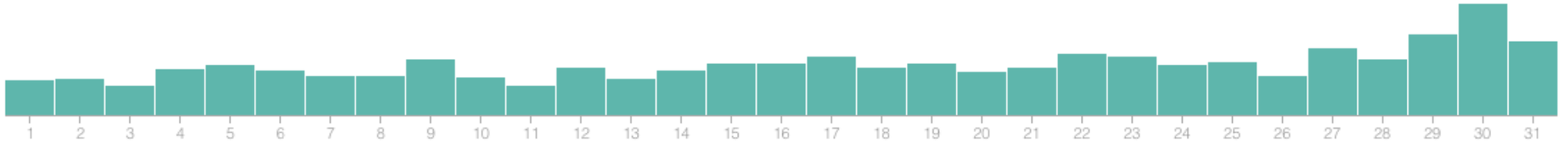
VALUE HISTOGRAM

MONTH

DAY OF WEEK



DAY OF MONTH



Sort: Default Edit

State
City
File_Number
CommitteeID
Occupation
Memo_Text
Report_Type
Trans_Date
Trans_Amount
Employer
ZIP
Name

Trans_Date

SUMMARY

Valid	1,537	100.0%
Unique	501	32.6%
Outliers	0	0.0%
Mismatched	0	0.0%
Missing	0	0.0%

STATISTICS

Minimum	2010-11-30 00:00:00
Lower Quartile	2011-11-22 00:00:00
Median	2012-05-24 00:00:00
Upper Quartile	2012-09-17 00:00:00
Maximum	2013-01-14 00:00:00
Average	2012-04-06 23:59:03

TOP VALUES

10172012	22
08312012	18
03302012	15
09302011	15
10092012	14
09272012	13
10222012	13
10232012	13
06302012	11
09292011	11
10152012	11
06292011	10
09182012	10
09242012	10

MISMATCHED VALUES

None

OUTLIERS

None

VALUE HISTOGRAM

Jan 2011 Apr 2011 Jul 2011 Oct 2011 Jan 2012 Apr 2012 Jul 2012 Oct 2012 Jan 2013

MONTH

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

DAY OF WEEK

Sun Mon Tue Wed Thu Fri Sat

DAY OF MONTH

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Overview

Sort: Default Edit

State

City

File_Number

CommitteeID

Occupation

Memo_Text

Report_Type

Trans_Date

Trans_Amount

Employer

ZIP

Name

Details

Trans_Date

SUMMARY

Valid	1,537	100.0%
Unique	501	32.6%
Outliers	0	0.0%
Mismatched	0	0.0%
Missing	0	0.0%

STATISTICS

Minimum	2010-11-30 00:00:00
Lower Quartile	2011-11-22 00:00:00
Median	2012-05-24 00:00:00
Upper Quartile	2012-09-17 00:00:00
Maximum	2013-01-14 00:00:00
Average	2012-04-06 23:59:03

TOP VALUES

10172012	22
08312012	18
03302012	15
09302011	15
10092012	14
09272012	13
10222012	13
10232012	13
06302012	11
09292011	11
10152012	11
06292011	10
09182012	10
09242012	10

MISMATCHED VALUES

None

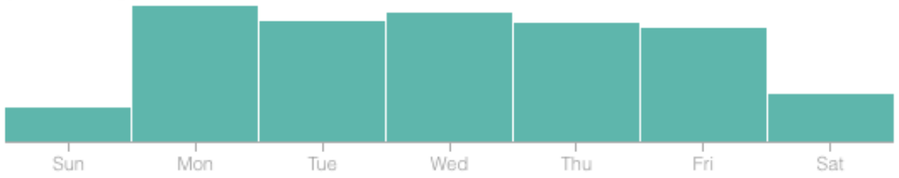
OUTLIERS

None

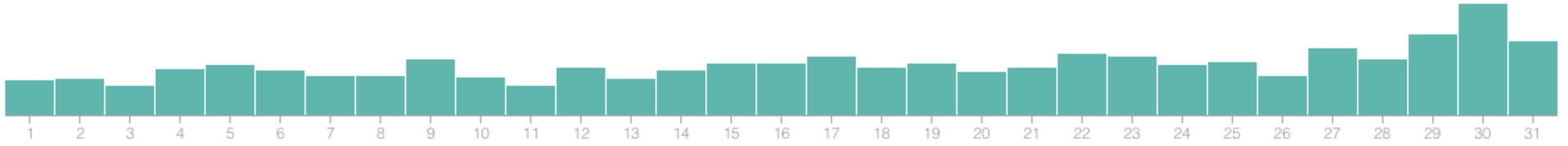
VALUE HISTOGRAM

MONTH

DAY OF WEEK



DAY OF MONTH



Effective

21 Columns 1,537 Rows 6 Data Types Column Details

Sort: Default Edit

State

ABC City

WASHINGTON 40 2.60%

File_Number

ABC CommitteeID

ABC Occupation

ABC Memo_Text

ABC Report_Type

Trans_Date

Trans_Amount

ABC Employer

State

21 Columns 1,537 Rows 6 Data Types Column Details

Sort: Default Edit

State

ABC City

WASHINGTON 40 2.60%

File_Number

ABC CommitteeID

ABC Occupation

ABC Memo_Text

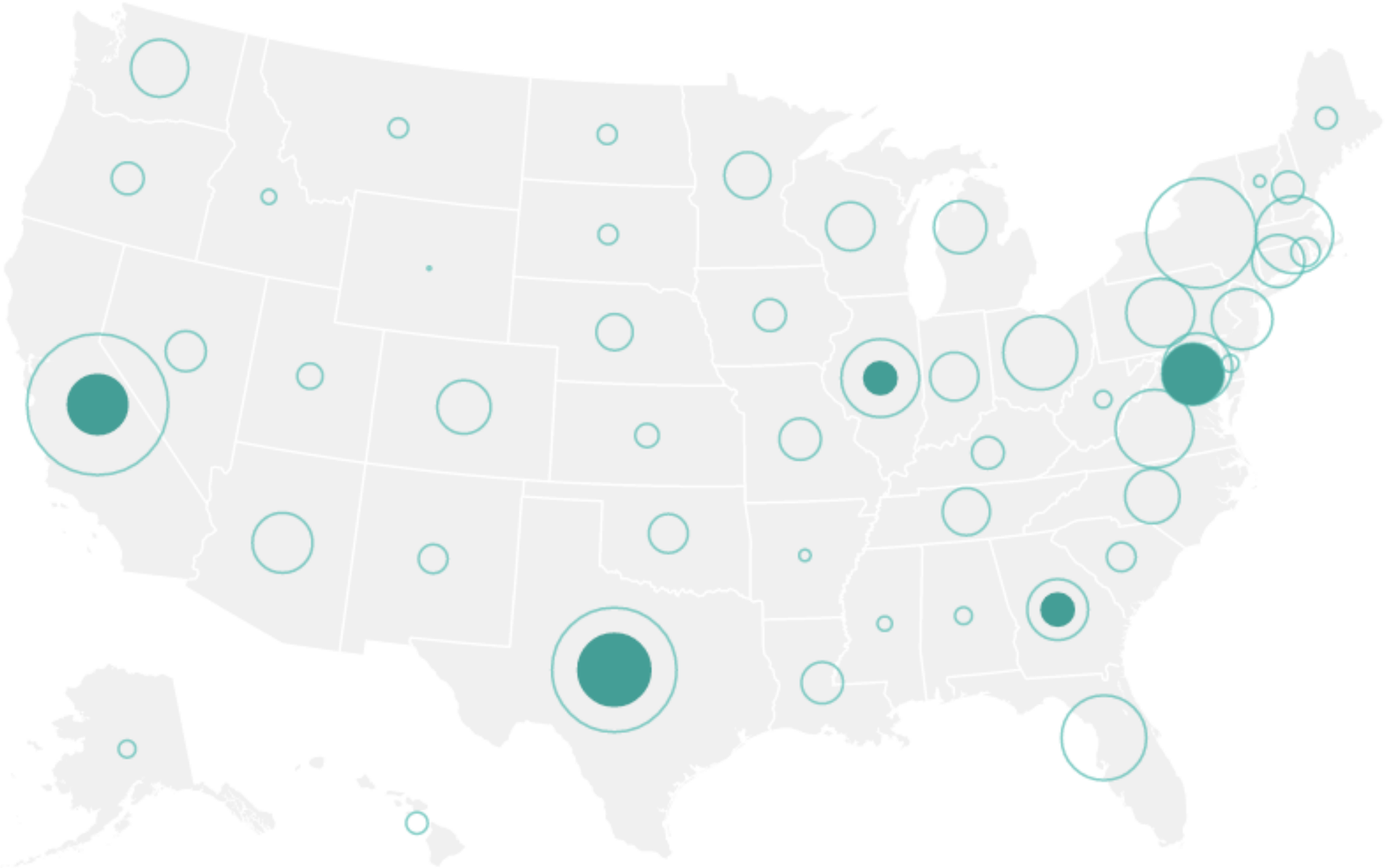
ABC Report_Type

Trans_Date

Trans_Amount

ABC Employer

State



Scalable

Summarize Random Sample

Trans_Amount

SUMMARY

Unique	212	8.9%
Valid	2,391	100.0%
Mismatched	0	0.0%
Missing	0	0.0%
Extreme	212	8.9%

FREQUENT VALUES

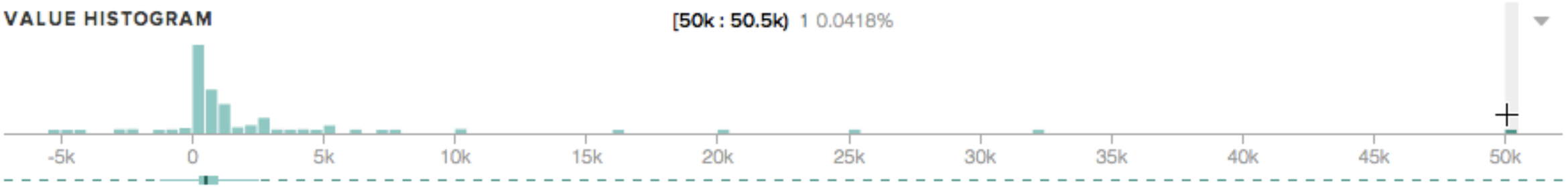
250	524
500	452
1,000	316
200	130
300	106

INVALID VALUES

None

VALUE HISTOGRAM

[50k : 50.5k) 1 0.0418%



FREQUENT VALUES

Summarize Full Data Set

Trans_Amount

SUMMARY

Unique	4,657	0.3%
Valid	1,539,514	100.0%
Mismatched	0	0.0%
Missing	0	0.0%
Extreme	148,056	9.6%

FREQUENT VALUES

250	339,126
500	278,466
1,000	204,290
200	100,182
2,600	82,535

INVALID VALUES

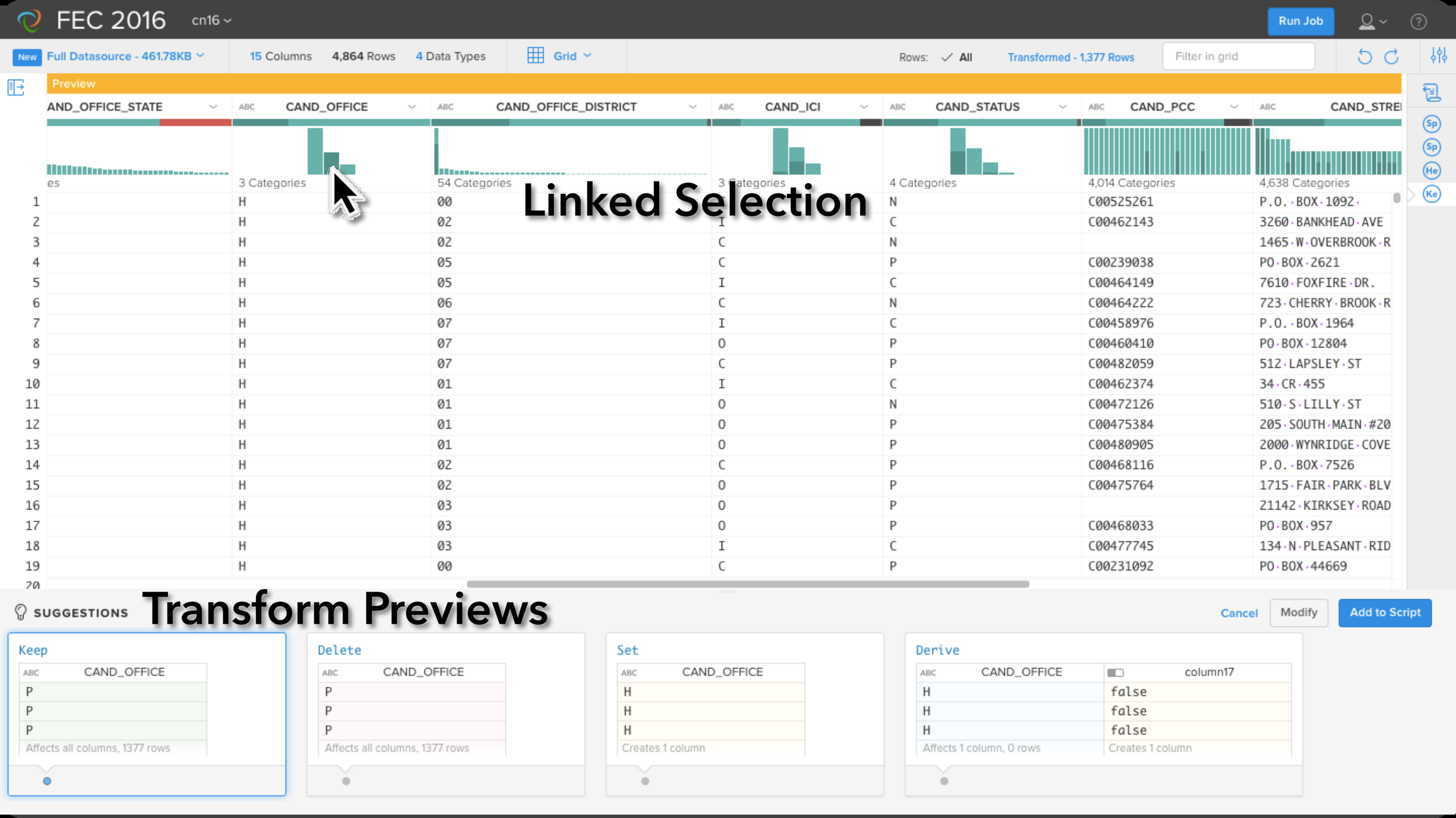
None

VALUE HISTOGRAM

[15M : 15.1M) 1 0.0001%

FREQUENT VALUES

Actionable


FEC 2016
cn16
Run Job
New
Full Datasource - 461.78KB
15 Columns
4,864 Rows
4 Data Types
Grid
Rows:
All
Transformed - 1,377 Rows
Filter in grid
Preview
AND_OFFICE_STATE
CAND_OFFICE
CAND_OFFICE_DISTRICT
CAND_ICI
CAND_STATUS
CAND_PCC
CAND_STREI
es
3 Categories
54 Categories
3 Categories
4 Categories
4,014 Categories
4,638 Categories
1 H 00 N C00525261 P.O. BOX 1092
2 H 02 I C C00462143 3260 BANKHEAD AVE
3 H 02 C N 1465 W OVERBROOK R
4 H 05 C P C00239038 PO BOX 2621
5 H 05 I C C00464149 7610 FOXFIRE DR.
6 H 06 C N C00464222 723 CHERRY BROOK R
7 H 07 I C C00458976 P.O. BOX 1964
8 H 07 0 P C00460410 PO BOX 12804
9 H 07 C P C00482059 512 LAPSLEY ST
10 H 01 I C C00462374 34 CR 455
11 H 01 0 N C00472126 510 S LILLY ST
12 H 01 0 P C00475384 205 SOUTH MAIN #20
13 H 01 0 P C00480905 2000 WYNRIDGE COVE
14 H 02 C P C00468116 P.O. BOX 7526
15 H 02 0 P C00475764 1715 FAIR PARK BLV
16 H 03 0 P 21142 KIRKSEY ROAD
17 H 03 0 P C00468033 PO BOX 957
18 H 03 I C C00477745 134 N PLEASANT RID
19 H 00 C P C00231092 PO BOX 44669
Linked Selection
SUGGESTIONS
Transform Previews
Cancel
Modify
Add to Script
Keep
CAND_OFFICE
P
P
P
Affects all columns, 1377 rows
Delete
CAND_OFFICE
P
P
P
Affects all columns, 1377 rows
Set
CAND_OFFICE
H
H
H
Creates 1 column
Derive
CAND_OFFICE
column17
H false
H false
H false
Affects 1 column, 0 rows
Creates 1 column

TABLEAU

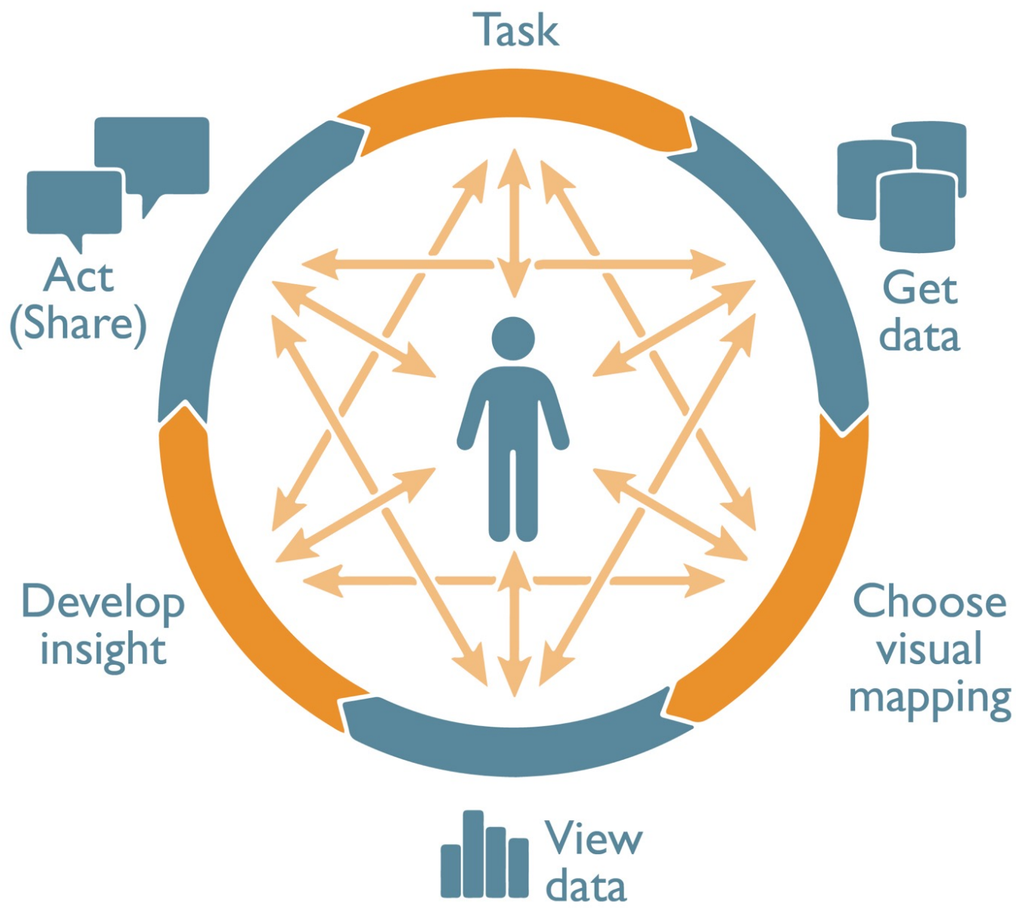
Task: Exploratory Analysis

Incremental

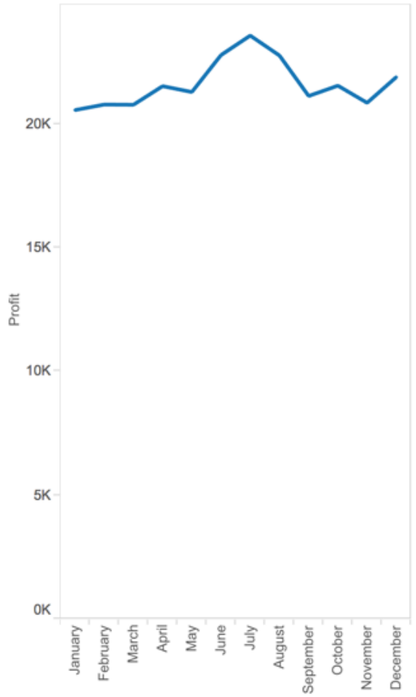
Expressive

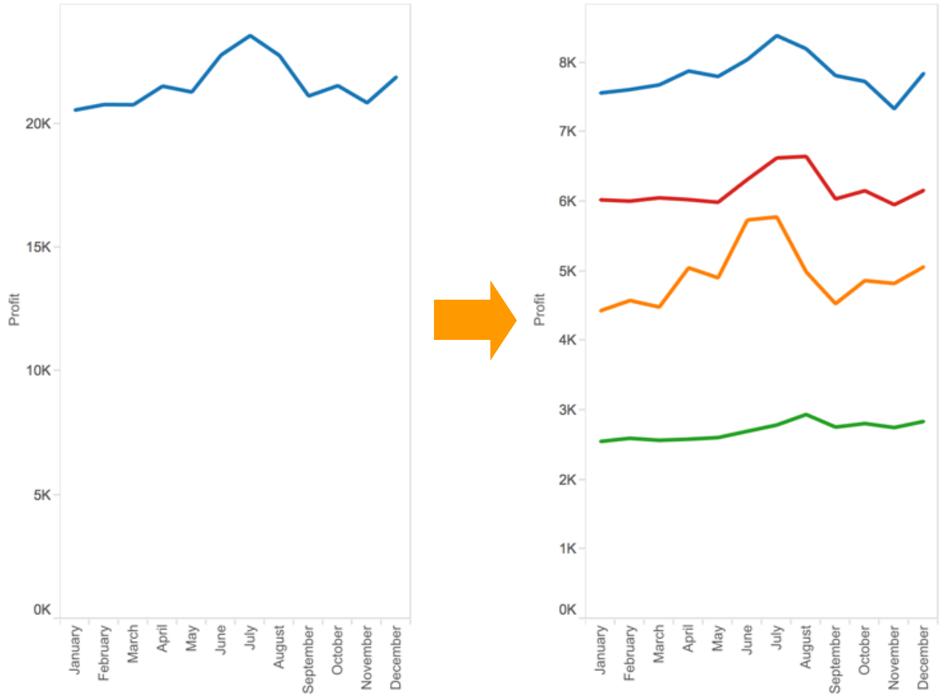
Unified

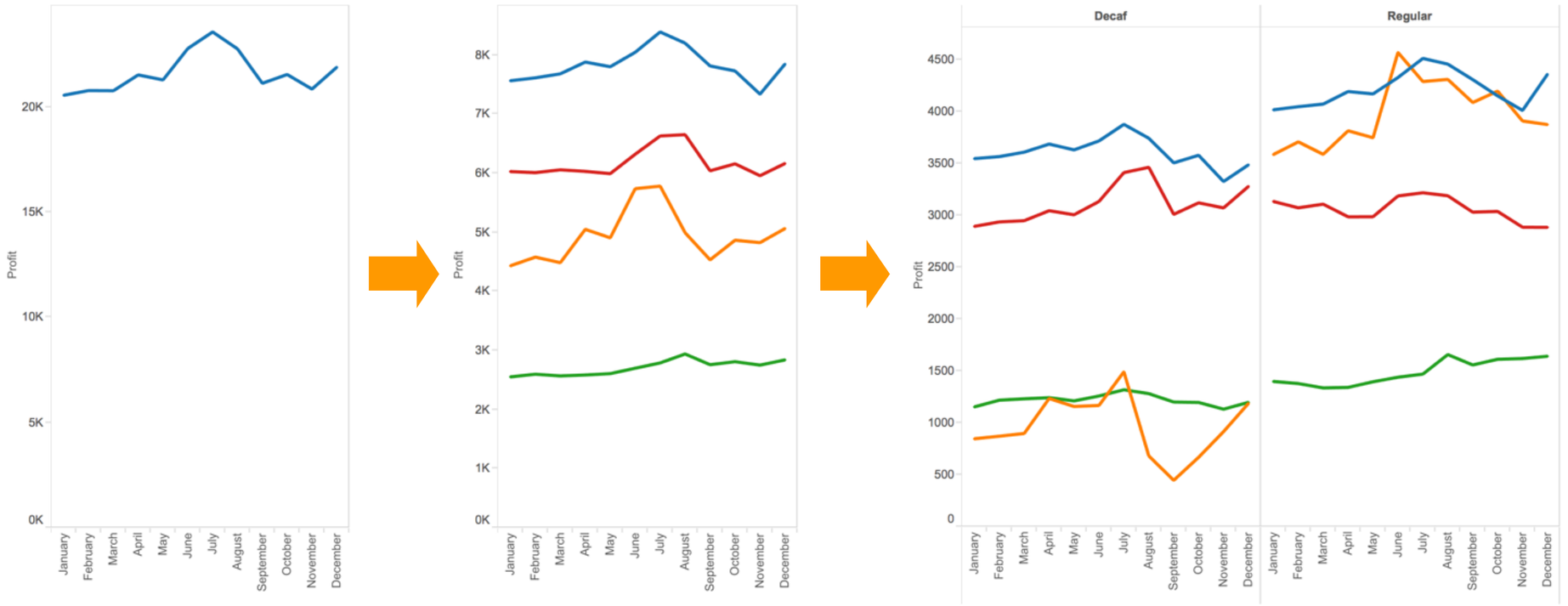
Direct



Incremental

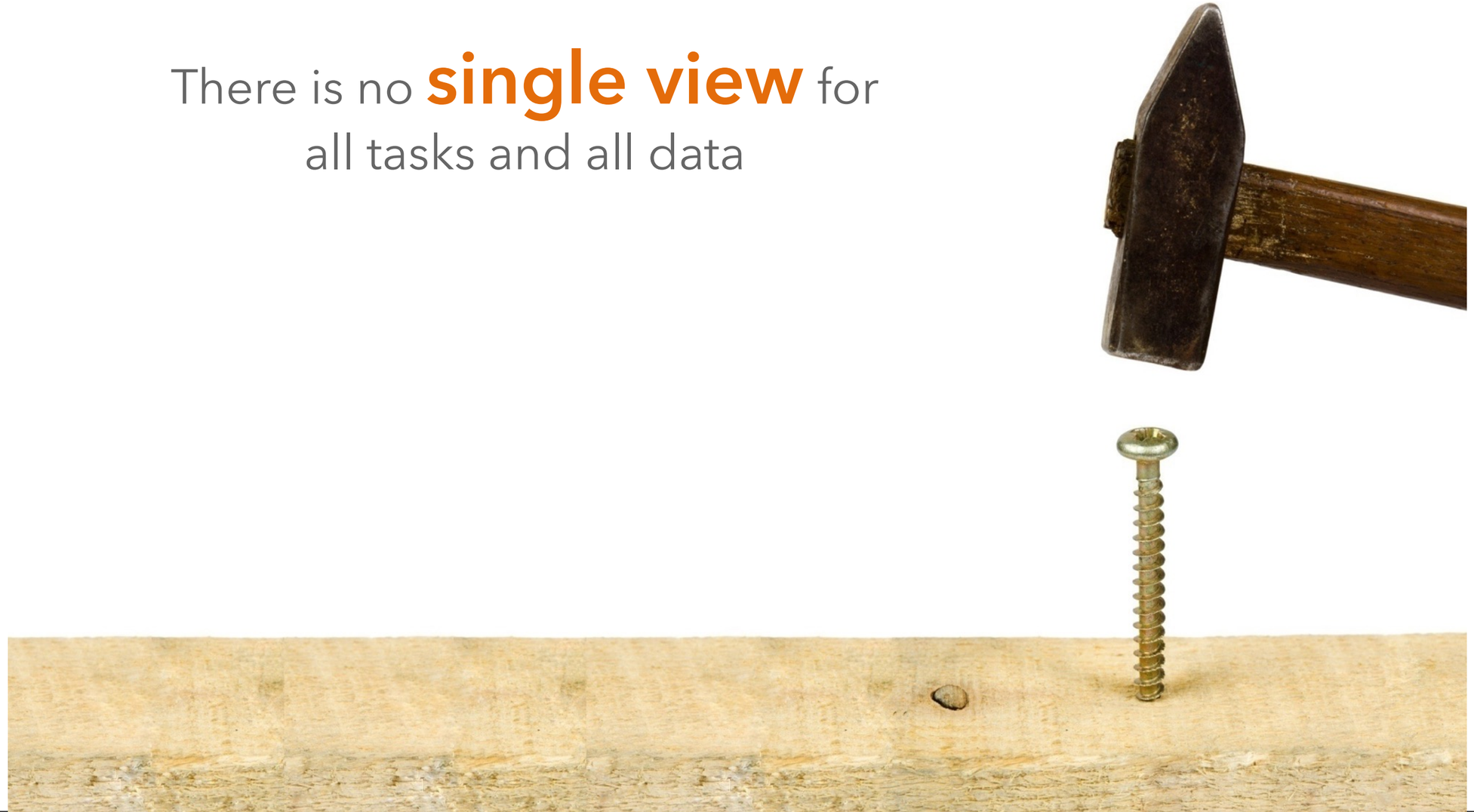

Profit
20K
15K
10K
5K
0K
January
February
March
April
May
June
July
August
September
October
November
December

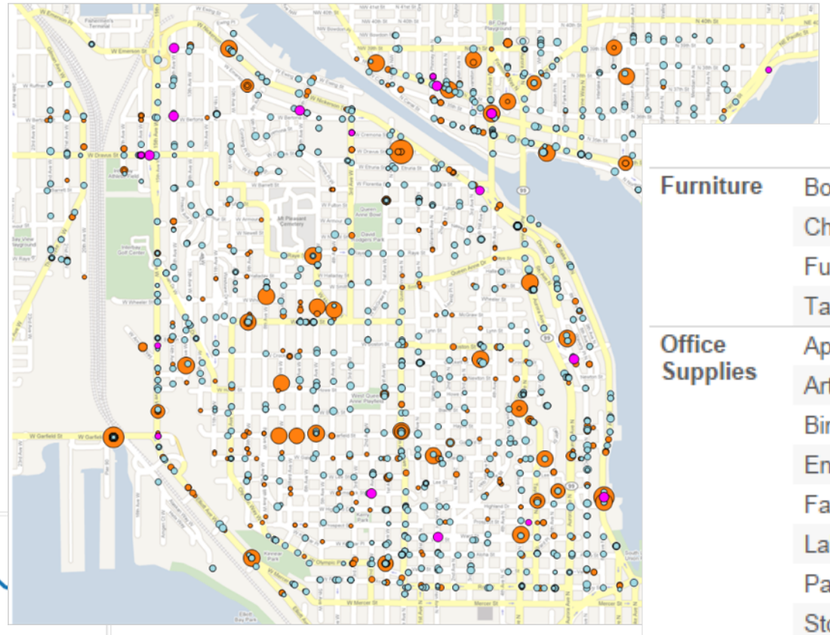

Profit
20K
15K
10K
5K
0K
January
February
March
April
May
June
July
August
September
October
November
December
Profit
8K
7K
6K
5K
4K
3K
2K
1K
0K
January
February
March
April
May
June
July
August
September
October
November
December

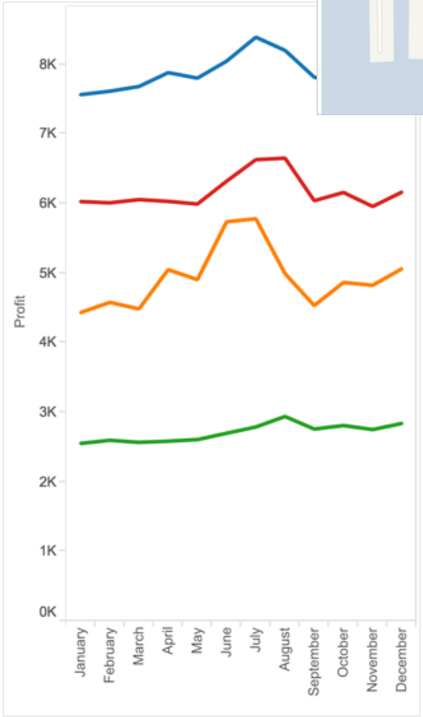

Profit
20K
15K
10K
5K
0K
January
February
March
April
May
June
July
August
September
October
November
December
Profit
8K
7K
6K
5K
4K
3K
2K
1K
0K
Decaf
Regular
Profit
4500
4000
3500
3000
2500
2000
1500
1000
500
0

Expressive

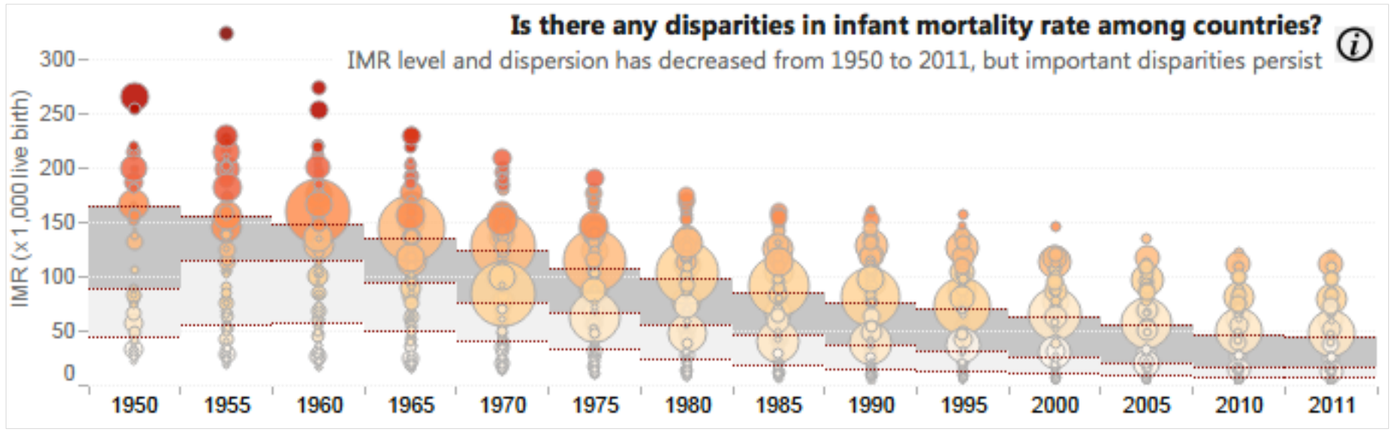
There is no **single view** for
all tasks and all data



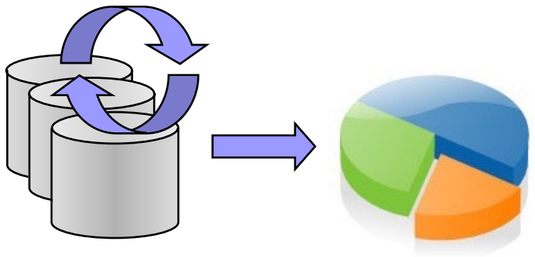




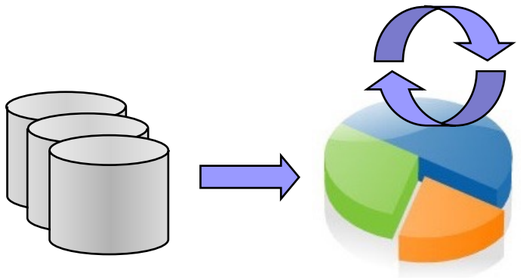
		Central	East	South	West
Furniture	Bookcases	-$1,998	-$1,168	$1,339	-$1,647
	Chairs	$6,593	$9,358	$6,612	$4,028
	Furnishings	-$3,906	$5,881	$3,443	$7,641
	Tables	-$3,560	-$11,025	-$4,623	$1,483
Office Supplies	Appliances	-$2,639	$8,391	$4,124	$8,261
	Art	$1,195	$1,900	$1,059	$2,374
	Binders	-$1,044	$11,268	$3,901	$16,097
	Envelopes	$1,778	$1,812	$1,465	$1,909
	Fasteners	$237	$264	$174	$275
	Labels	$1,073	$1,129	$1,041	$2,303
	Paper	$6,972	$9,015	$5,947	$12,119
	Storage	$1,970	$8,389	$2,274	$8,645
	Supplies	-$662	-$1,155	$2	$626
Technology	Accessories	$7,252	$11,196	$7,005	$16,485
	Copiers	$15,609	$17,023	$3,659	$19,327
	Machines	-$1,486	$6,929	-$1,439	-$619
	Phones	$12,323	$12,315	$10,767	$9,111



Unified

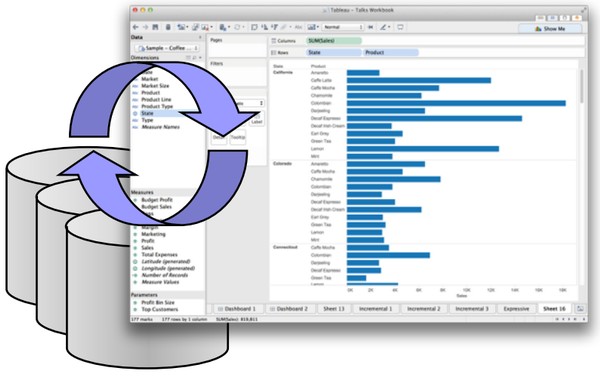


Traditional Reporting Tools



Traditional Visualization Tools

versus



Tableau

Direct

Sales by Product Type
Market
Central
East
South
West
Date
Product Type
Q1
Q2
Q3
Q4
Coffee
Espresso
Herbal Tea
Tea
0K
20K
40K
60K
Sales

Designing for Purpose

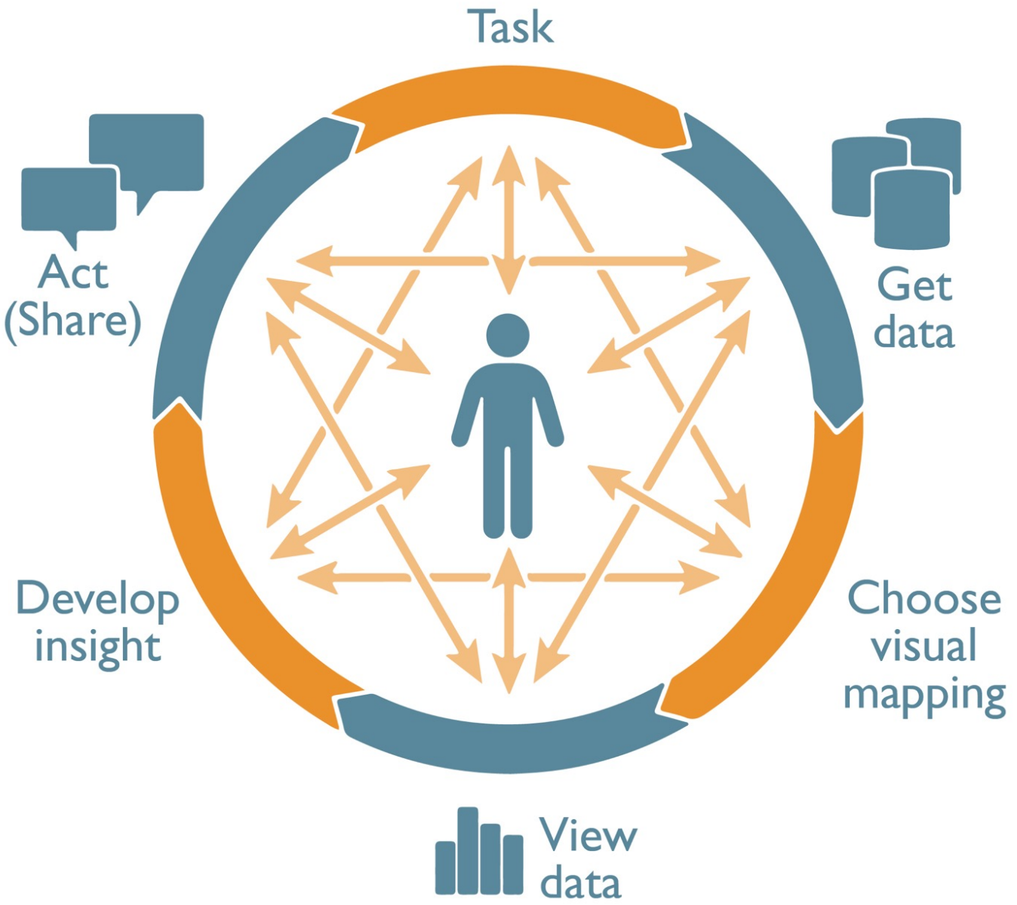
Tableau

Incremental

Expressive

Unified

Direct



Trifacta

Automatic

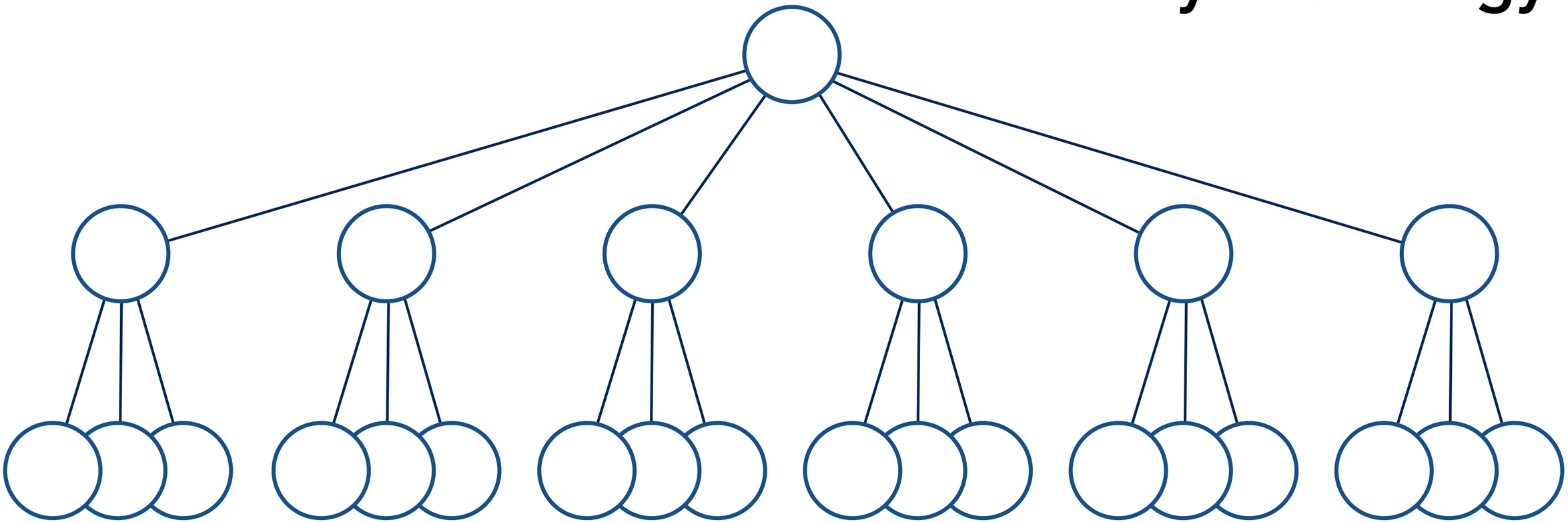
Effective

Scalable

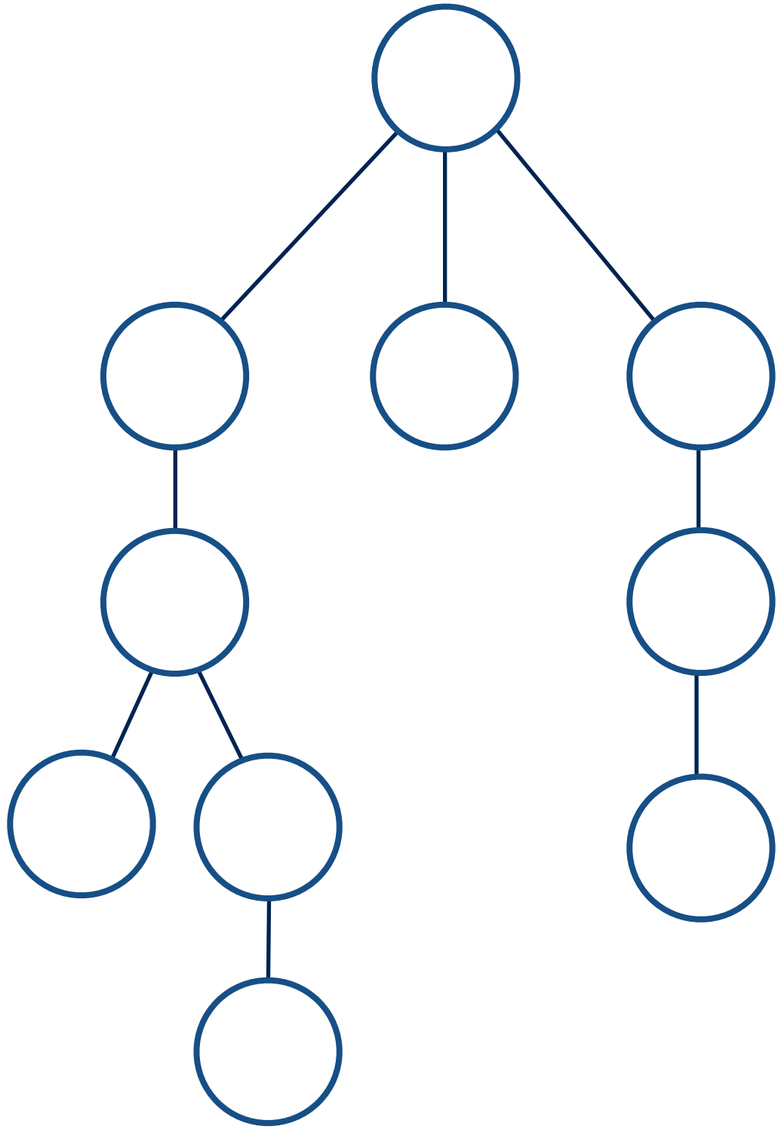
Actionable

Trifacta

Breadth-First Analysis Strategy



Tableau



Depth-First Analysis Strategy

How might we combine and extend these analysis strategies?

Data Voyager

vega.github.io/voyager/

Data Voyager

Bookmarks (0) Undo Redo

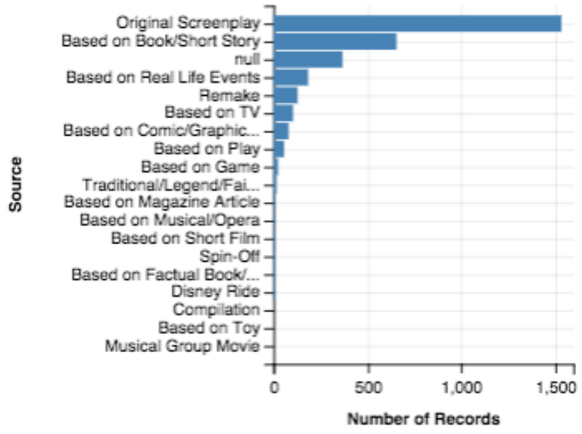
DATA

Movies

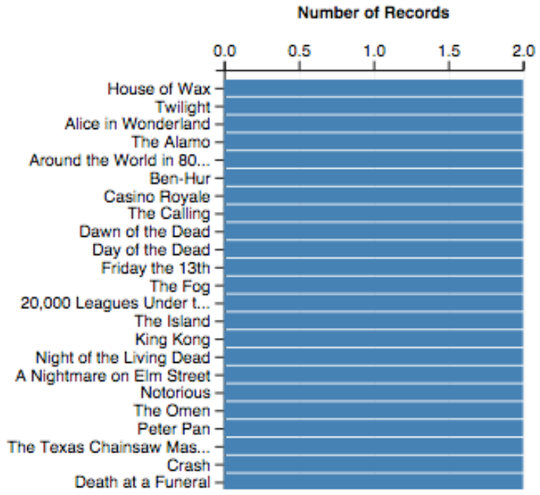
- Creative Type
- Director
- Distributor
- Major Genre
- MPAA Rating
- Release Date
- Source
- Title
- AUTO (IMDB Rating)
- AUTO (IMDB Votes)
- AUTO (Production Budget)
- AUTO (Rotten Tomatoes)
- AUTO (Running Time min)
- AUTO (US DVD Sales)
- AUTO (US Gross)
- AUTO (Worldwide Gross)
- COUNT

Reset

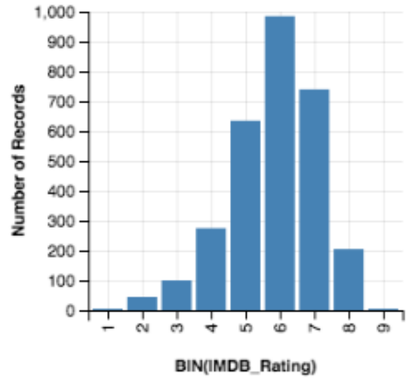
Source COUNT



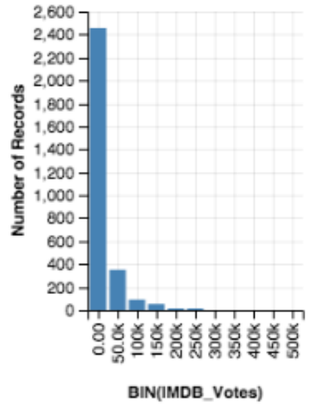
Title COUNT



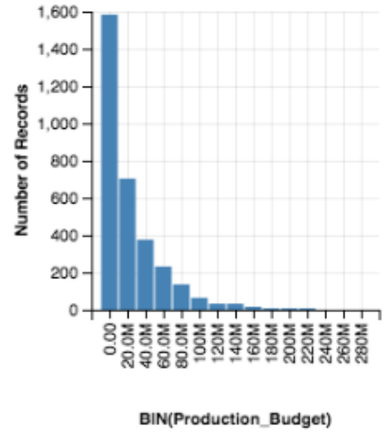
BIN (IMDB Rating) COUNT



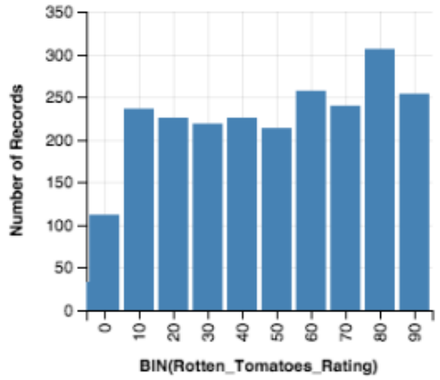
BIN (IMDB Votes) COUNT



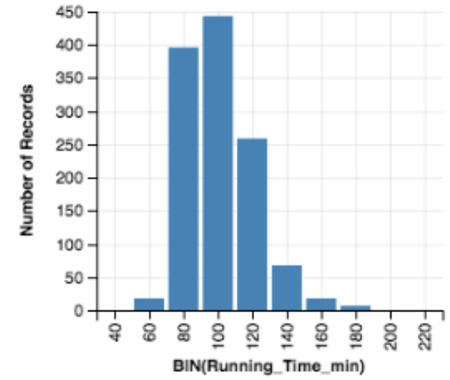
BIN (Production Budget) COUNT



BIN (Rotten Tomatoes Rating) COUNT



BIN (Running Time min) COUNT



BIN (US DVD Sales) COUNT

BIN (US Gross) COUNT

Data Voyager

vega.github.io/voyager/

Data Voyager

Bookmarks (0) Undo Redo

DATA

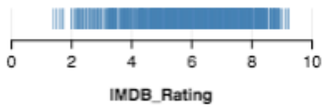
Movies

- Creative Type
- Director
- Distributor
- Major Genre
- MPAA Rating
- Release Date
- Source
- Title
- AUTO (IMDB Rating)
- AUTO (IMDB Votes)
- AUTO (Production Budget)
- AUTO (Rotten Tomatoes)
- AUTO (Running Time min)
- AUTO (US DVD Sales)
- AUTO (US Gross)
- AUTO (Worldwide Gross)
- COUNT

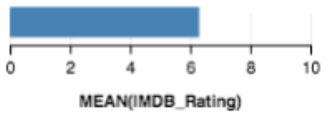
Reset

Showing Data Variations for # AUTO (IMDB Rating)

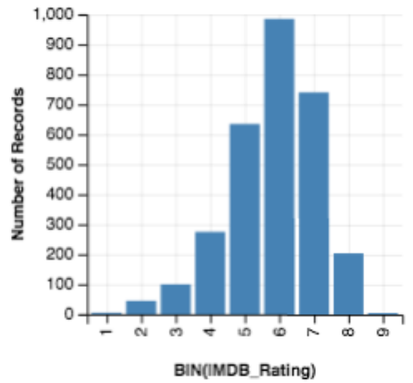
IMDB Rating



MEAN (IMDB Rating)

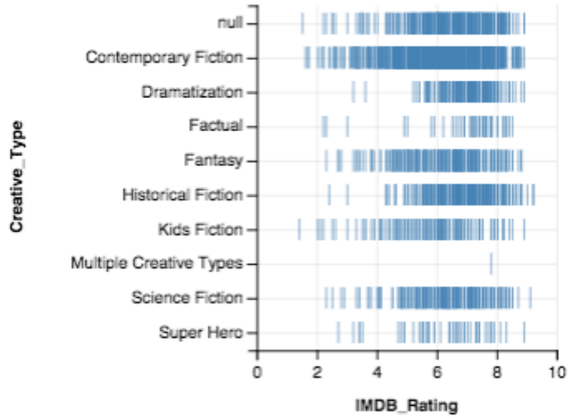


BIN (IMDB Rating) # COUNT

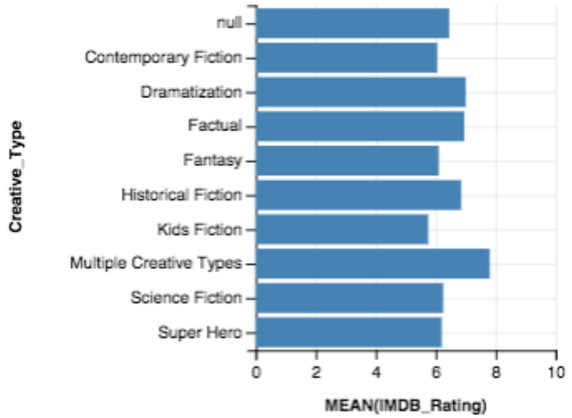


Showing Data Variations for # AUTO (IMDB Rating) **with one additional field.**

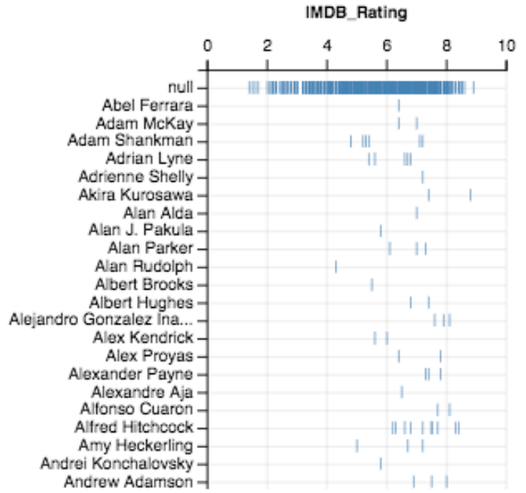
IMDB Rating Abc Creative Type



MEAN (IMDB Rating) Abc Creative Type



IMDB Rating Abc Director



MEAN (IMDB Rating) Abc Director

IMDB Rating Abc Distributor

MEAN (IMDB Rating) Abc Distributor

Data Voyager

vega.github.io/voyager/

Data Voyager

Bookmarks (0) Undo Redo

DATA

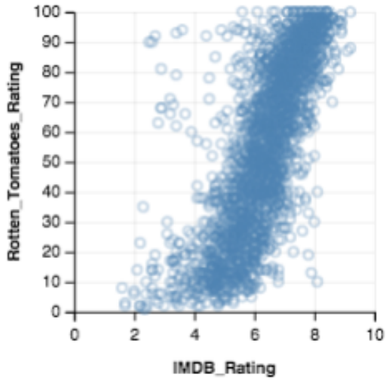
Movies

- Creative Type
- Director
- Distributor
- Major Genre
- MPAA Rating
- Release Date
- Source
- Title
- AUTO (IMDB Rating)
- AUTO (IMDB Votes)
- AUTO (Production Budget)
- AUTO (Rotten Tomatoes)
- AUTO (Running Time min)
- AUTO (US DVD Sales)
- AUTO (US Gross)
- AUTO (Worldwide Gross)
- COUNT

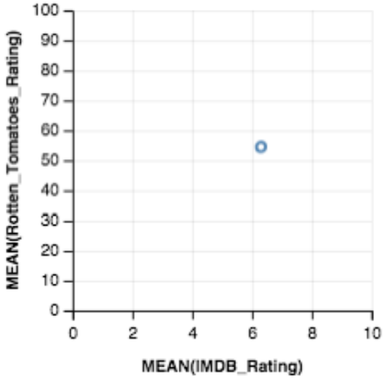
Reset

Showing Data Variations for # AUTO (IMDB Rating) # AUTO (Rotten Tomatoes Rating)

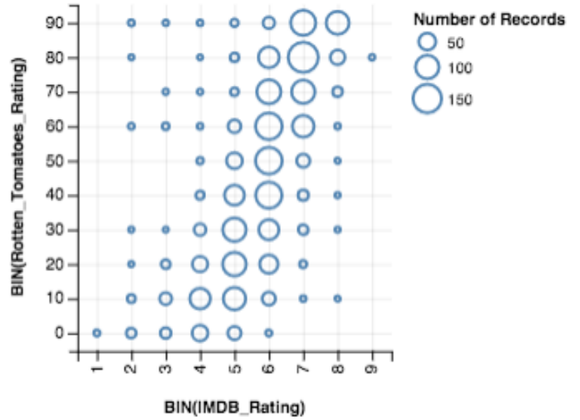
IMDB Rating # Rotten Tomatoes Rating



MEAN (IMDB Rating) # MEAN (Rotten Tomatoes Rating)

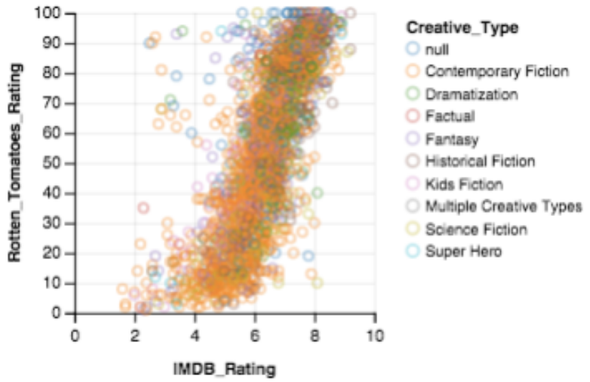


BIN (IMDB Rating) # BIN (Rotten Tomatoes Rating) # COUNT

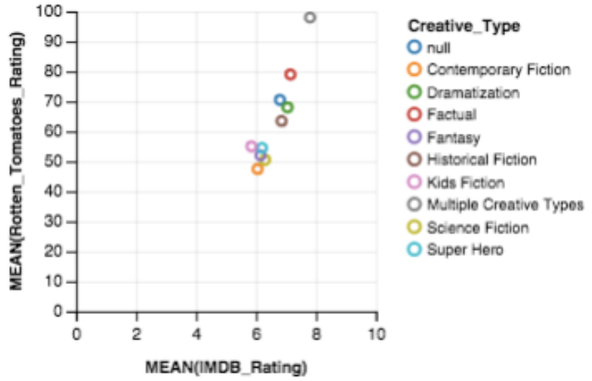


Showing Data Variations for # AUTO (IMDB Rating) # AUTO (Rotten Tomatoes Rating) **with one additional field.**

IMDB Rating # Rotten Tomatoes Rating Abc Creative Type



MEAN (IMDB Rating) # MEAN (Rotten Tomatoes Rating) Abc Creative Type



IMDB Rating # Rotten Tomatoes Rating Abc Director

MEAN (IMDB Rating) # MEAN (Rotten Tomatoes Rating) Abc Director

IMDB Rating # Rotten Tomatoes Rating Abc Distributor

Data Voyager

vega.github.io/voyager/

Data Voyager

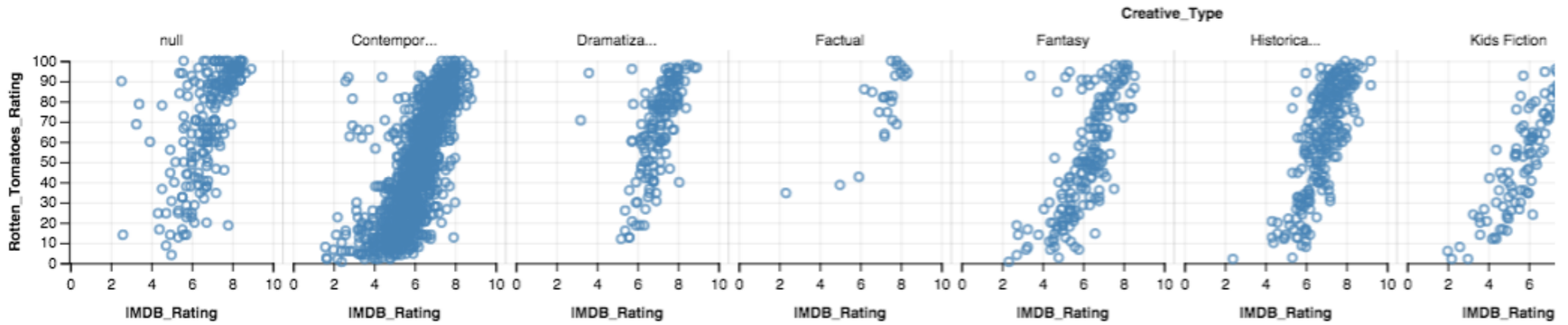
Bookmarks (0) Undo Redo

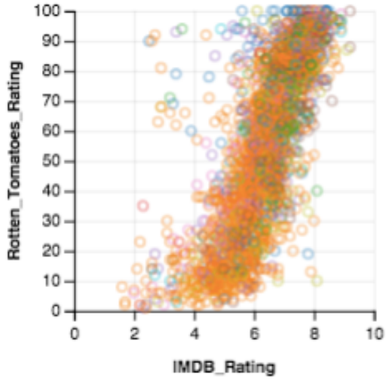
DATA

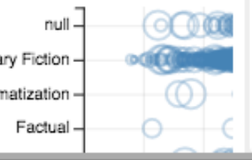
Various Visualizations Showing # IMDB Rating # Rotten Tomatoes Rating Abc Creative Type

Close

LOG X LOG Y







User

Data Set
Voyager
Visualization Browser
User

User

Compass

Recommendation Engine

Data Schema
& Statistics

Voyager

Visualization Browser

1. Select **data variables**
2. Apply **transformations**
3. Pick visual **encodings**

User

Compass
Recommendation Engine

Data Schema
& Statistics

Voyager
Visualization Browser

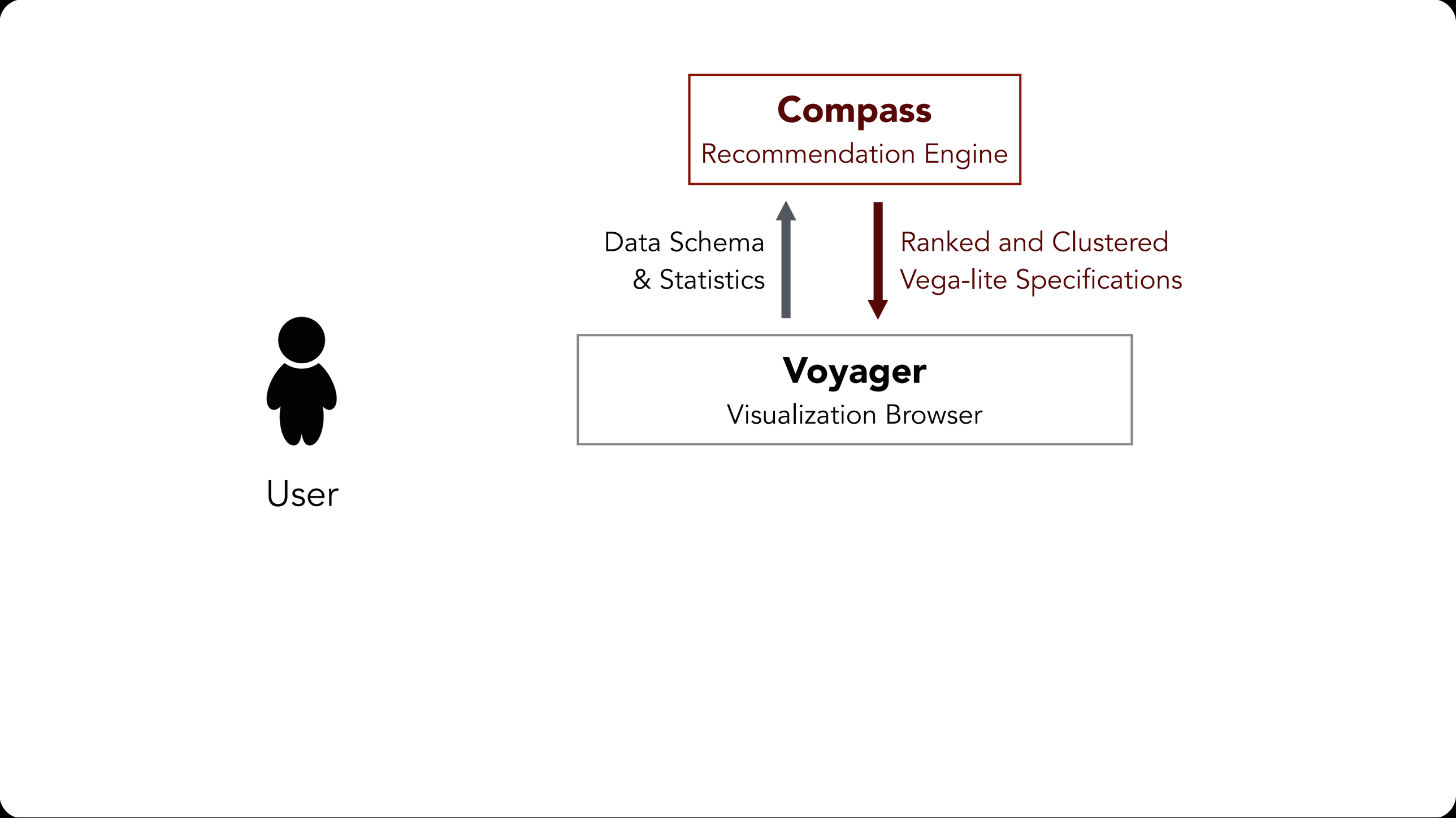
Constrain & rank choices by **data type**, **statistics** & **perceptual principles.**

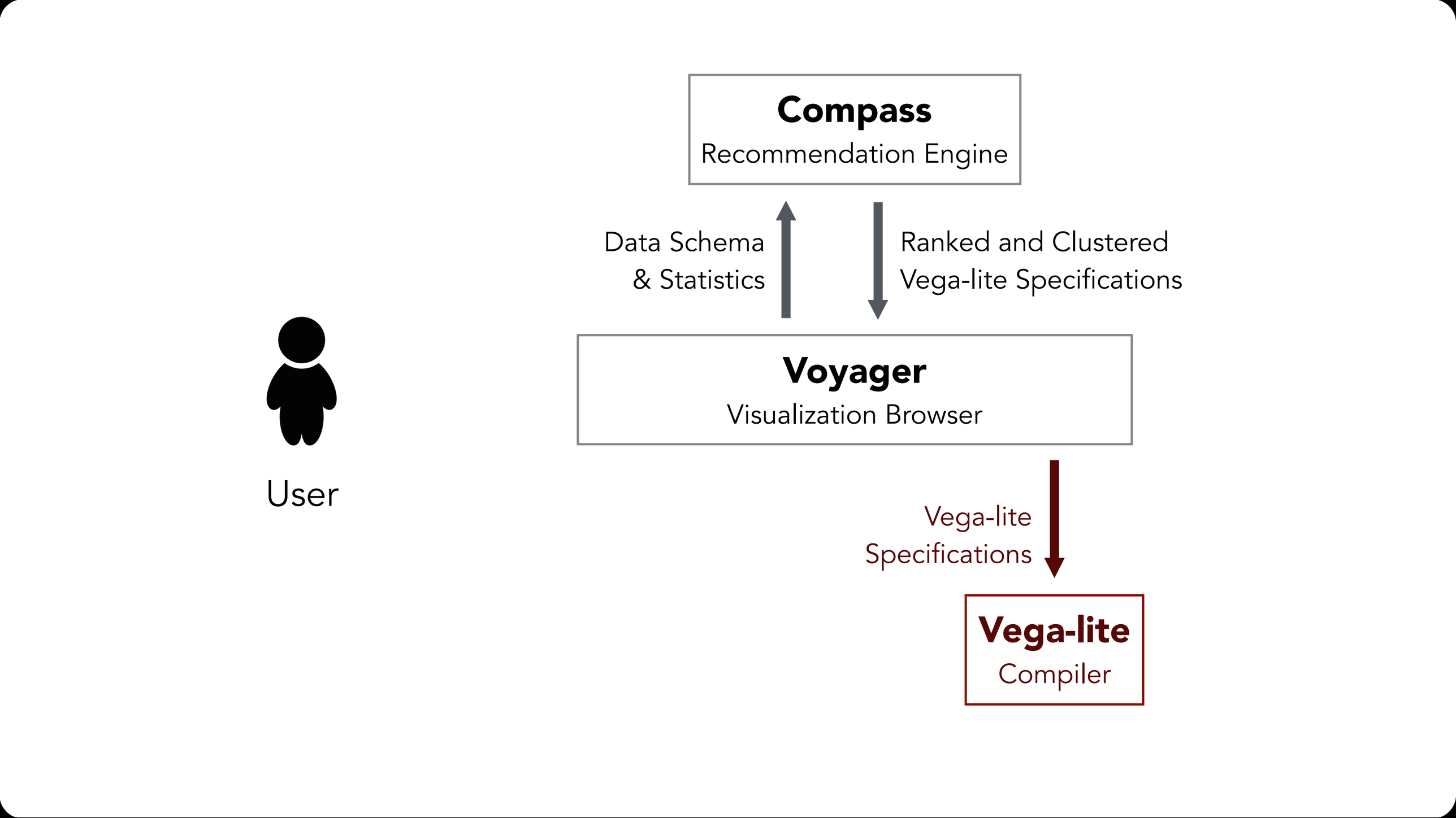
User

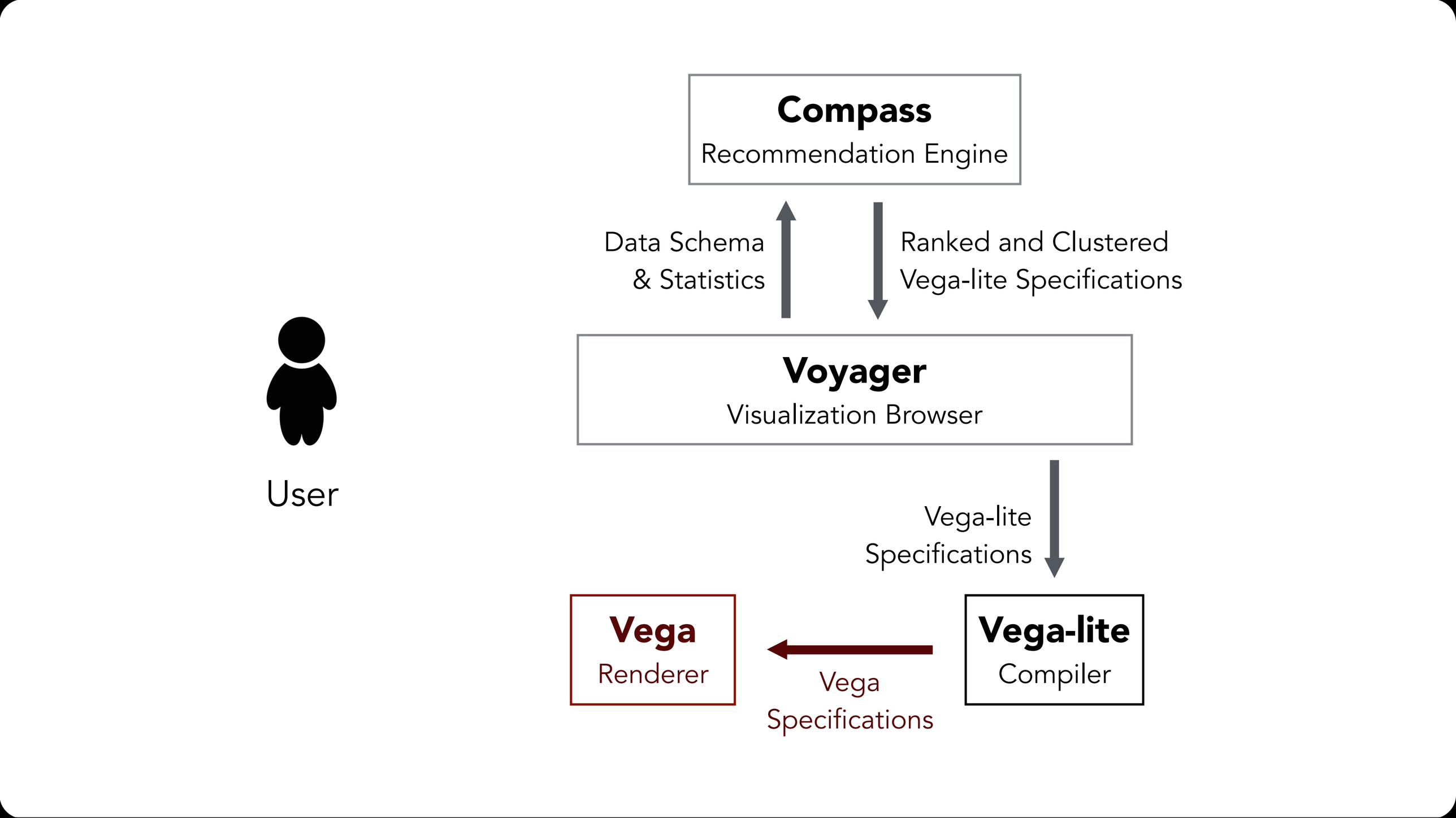
Compass
Recommendation Engine

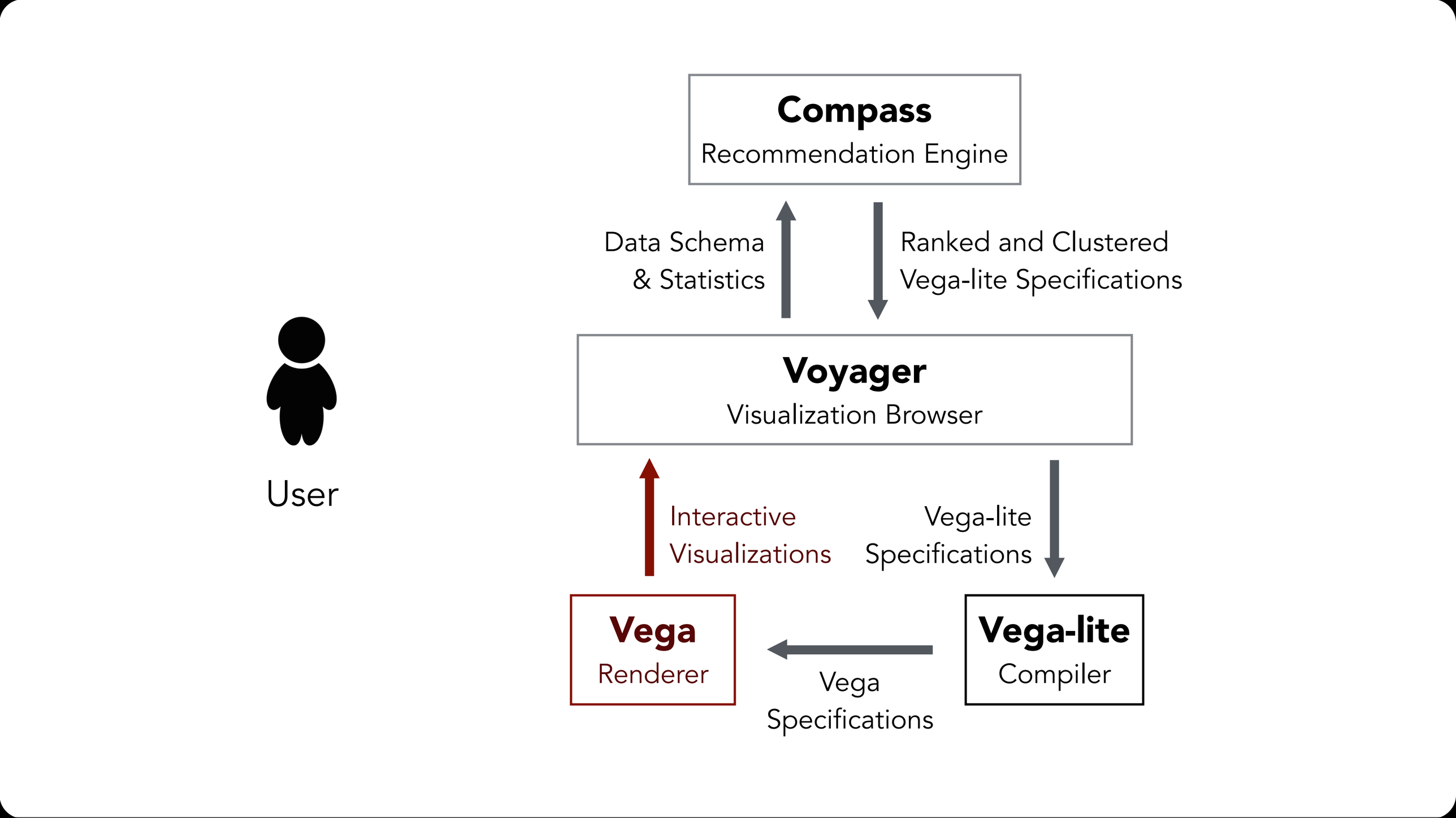
Data Schema & Statistics

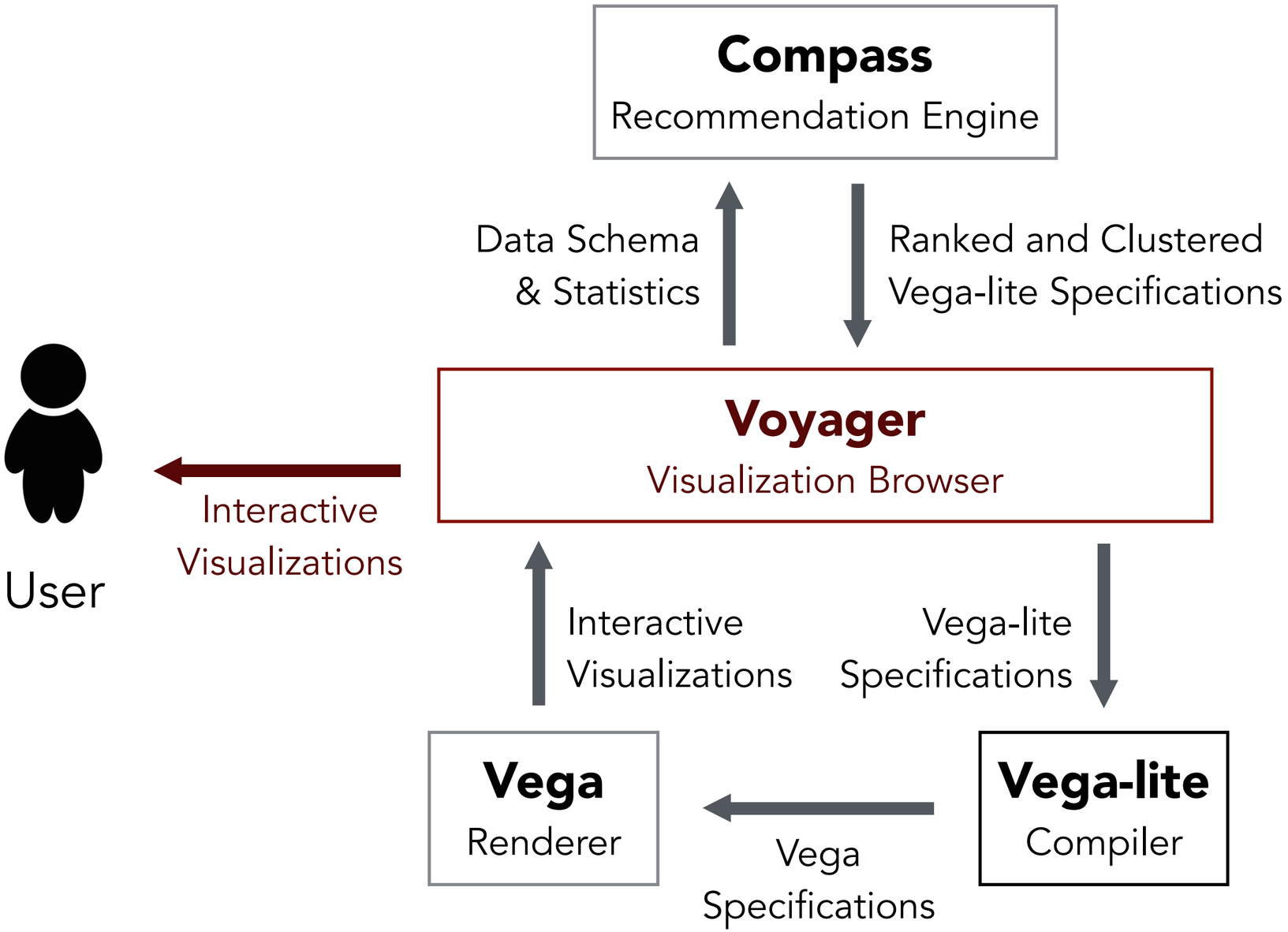
Voyager
Visualization Browser

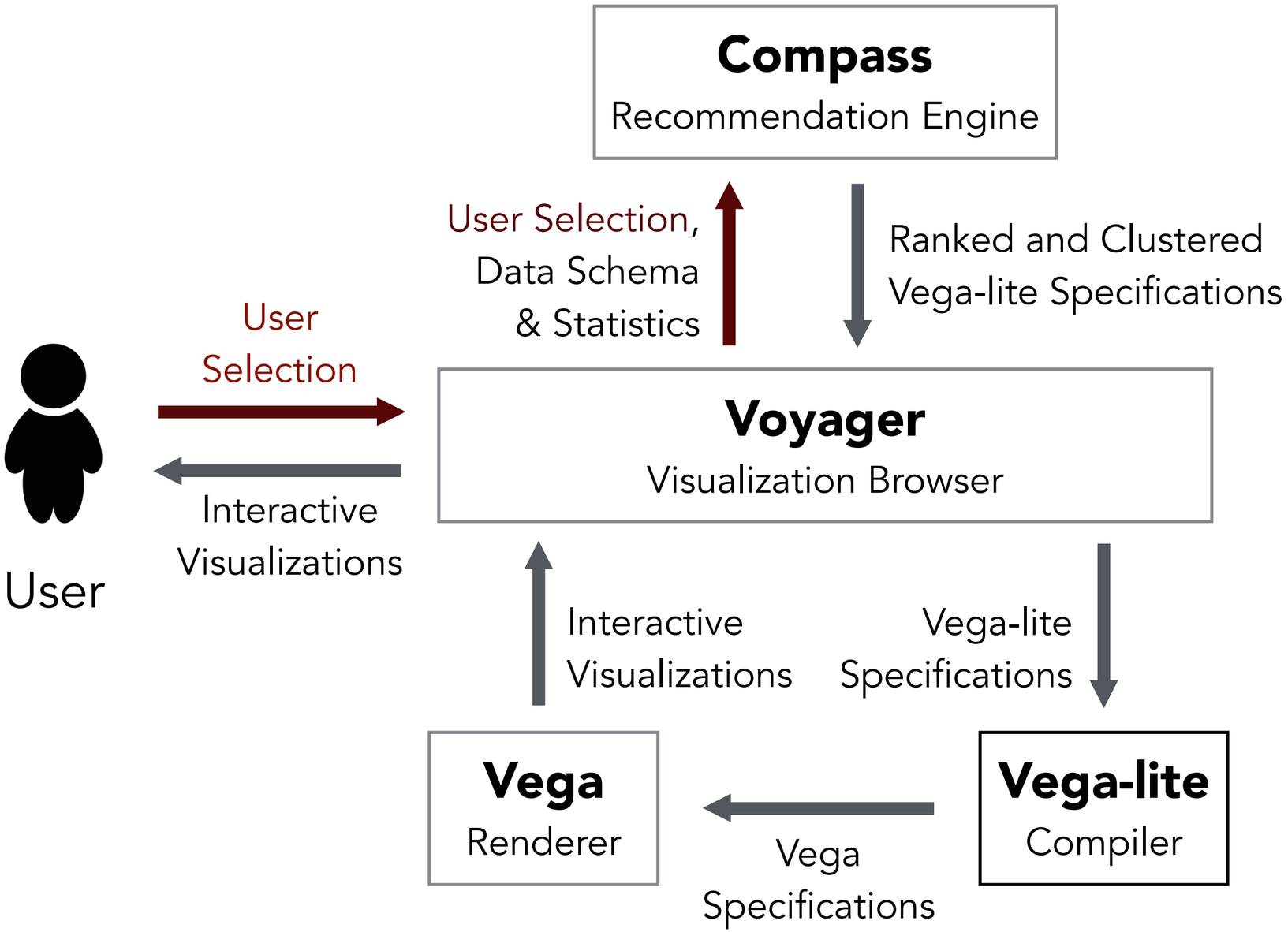

Compass
Recommendation Engine
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
Voyager
Visualization Browser
User


Compass
Recommendation Engine
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
Voyager
Visualization Browser
User
Vega-lite
Specifications
Vega-lite
Compiler


Compass
Recommendation Engine
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
Voyager
Visualization Browser
User
Vega-lite
Specifications
Vega
Renderer
Vega
Specifications
Vega-lite
Compiler


Compass
Recommendation Engine
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
Voyager
Visualization Browser
User
Interactive
Visualizations
Vega-lite
Specifications
Vega
Renderer
Vega
Specifications
Vega-lite
Compiler

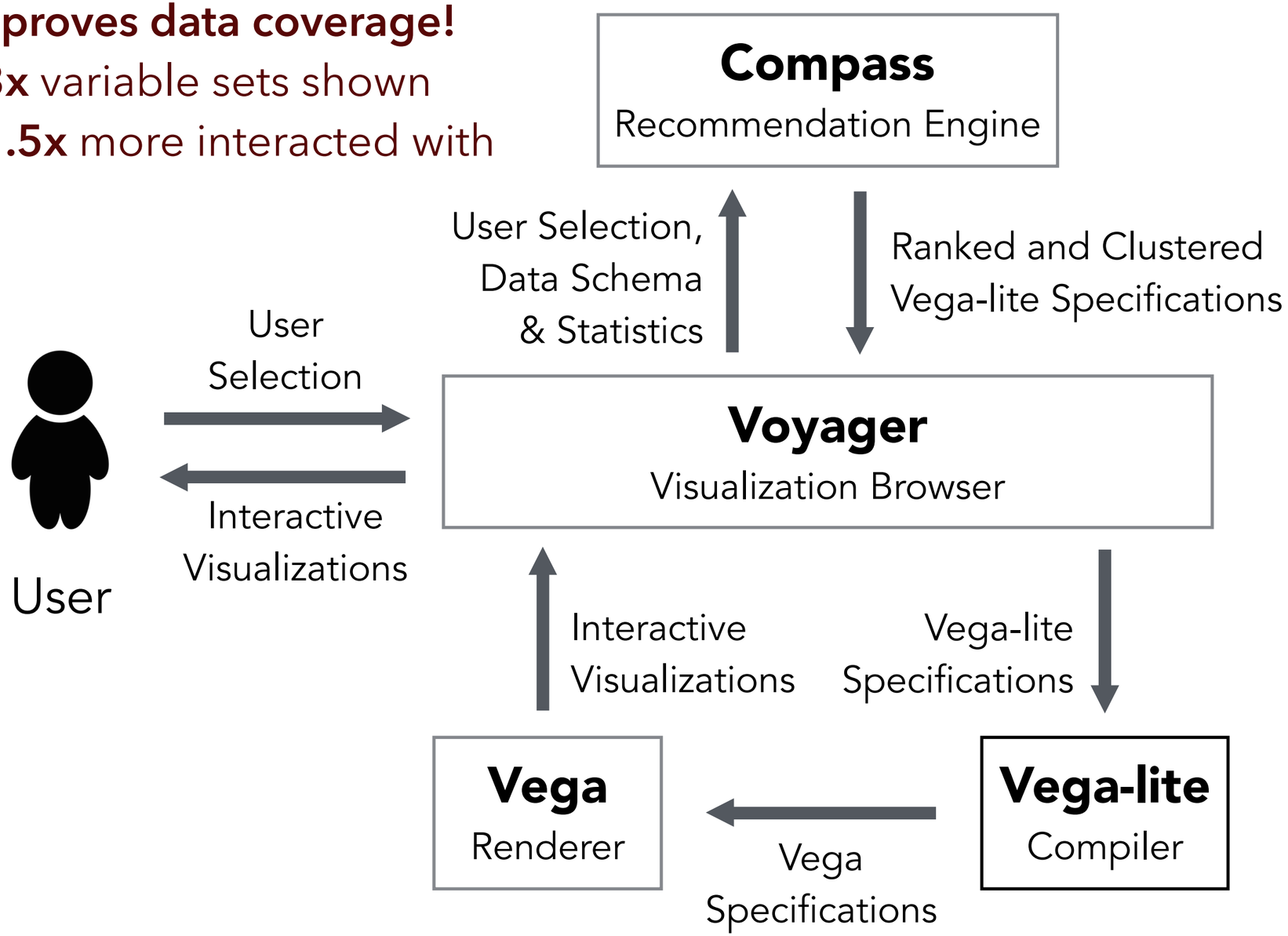

Compass
Recommendation Engine
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
Voyager
Visualization Browser
User
Interactive
Visualizations
Interactive
Visualizations
Vega-lite
Specifications
Vega
Renderer
Vega-lite
Compiler
Vega
Specifications


Compass
Recommendation Engine
User Selection,
Data Schema
& Statistics
Ranked and Clustered
Vega-lite Specifications
User
Selection
Interactive
Visualizations
User
Voyager
Visualization Browser
Interactive
Visualizations
Vega-lite
Specifications
Vega
Renderer
Vega
Specifications
Vega-lite
Compiler

Improves data coverage!

+**3x** variable sets shown

+**1.5x** more interacted with

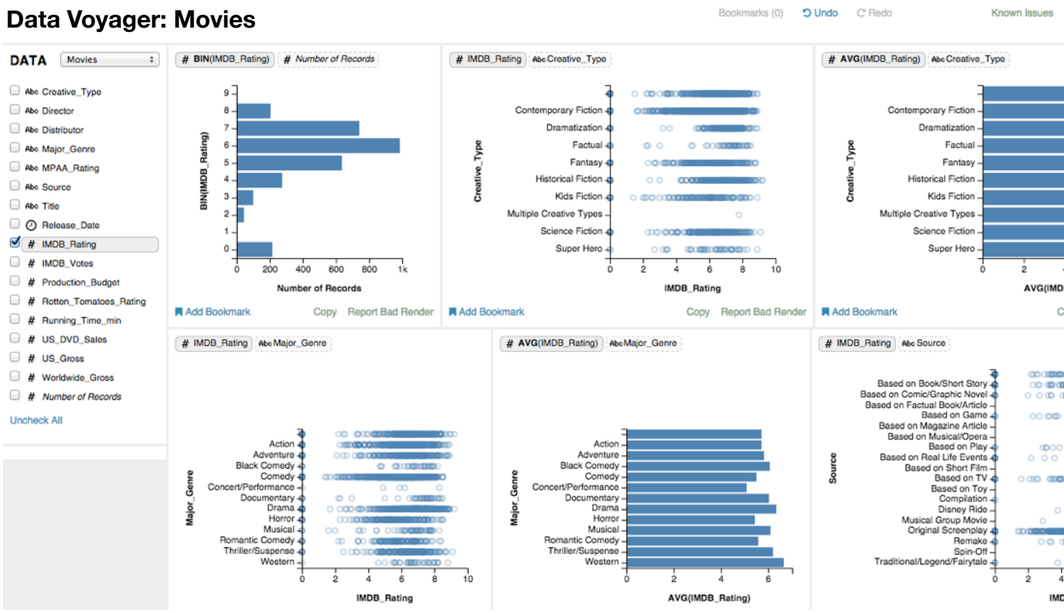


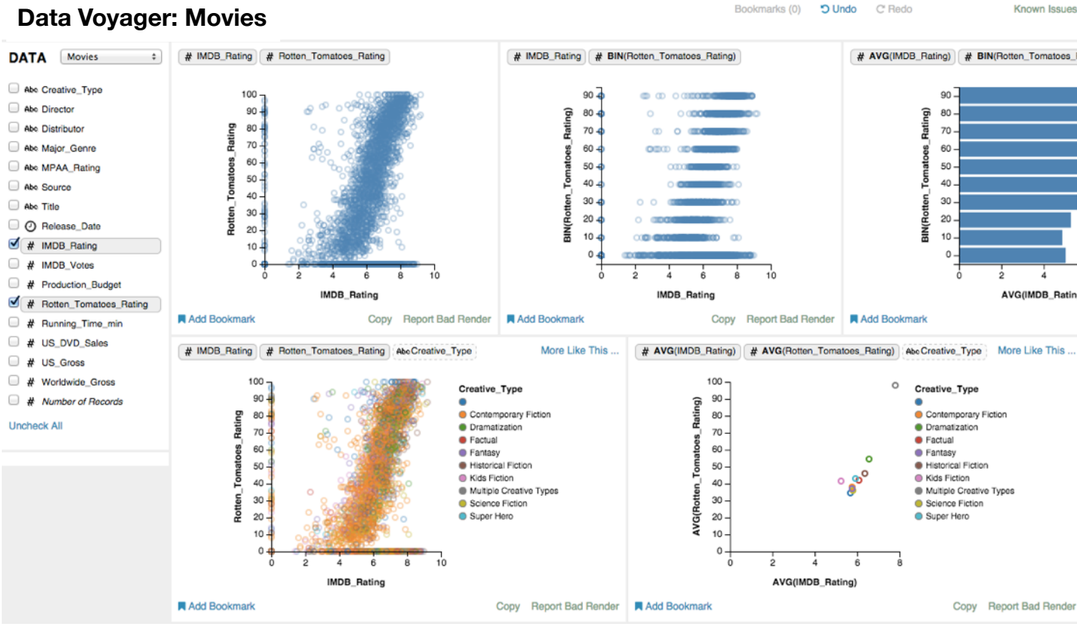
User Feedback

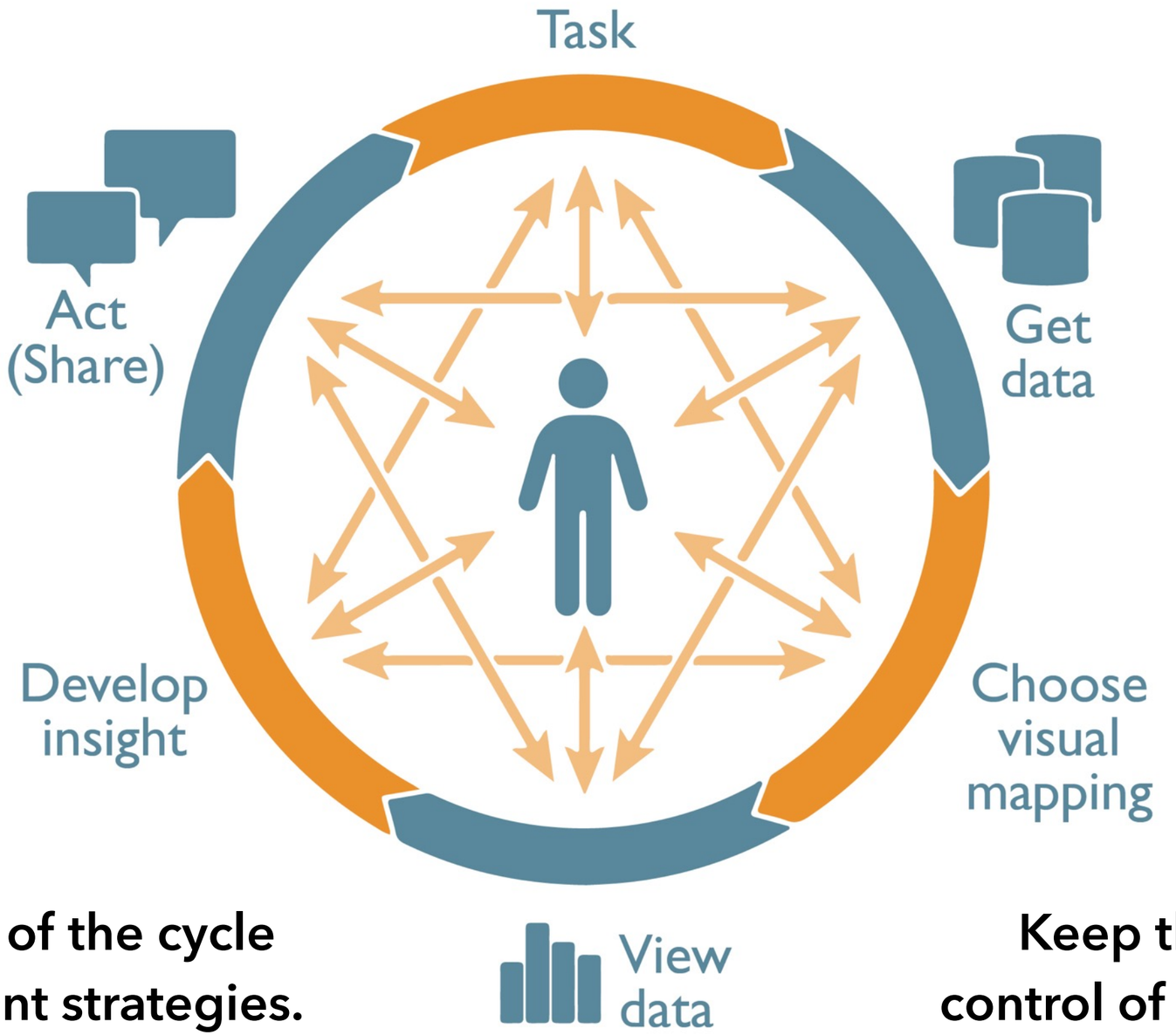
"Voyager gave me a lot of options I wouldn't have thought about on my own, it encouraged me to look more deeply at data, even data I didn't know a lot about."

Voyager *"prompted me to explore new questions in a way that didn't derail me from answering follow-up questions."*

"I would start with Voyager but want to go and switch to [Tableau] to dive into my question. Once that question was answered, I would like to switch back to Voyager."







Different iterations of the cycle may require different strategies.

Keep the analyst in control of the process!

From Profiling to Analysis

DESIGNING VISUALIZATION TOOLS FOR PURPOSE

Jeffrey Heer
Trifacta / U. Washington

Jock Mackinlay
Tableau Software