Predictive Interaction

Jeffrey Heer @jeffrey_heer
U. Washington / Trifacta

prefuse

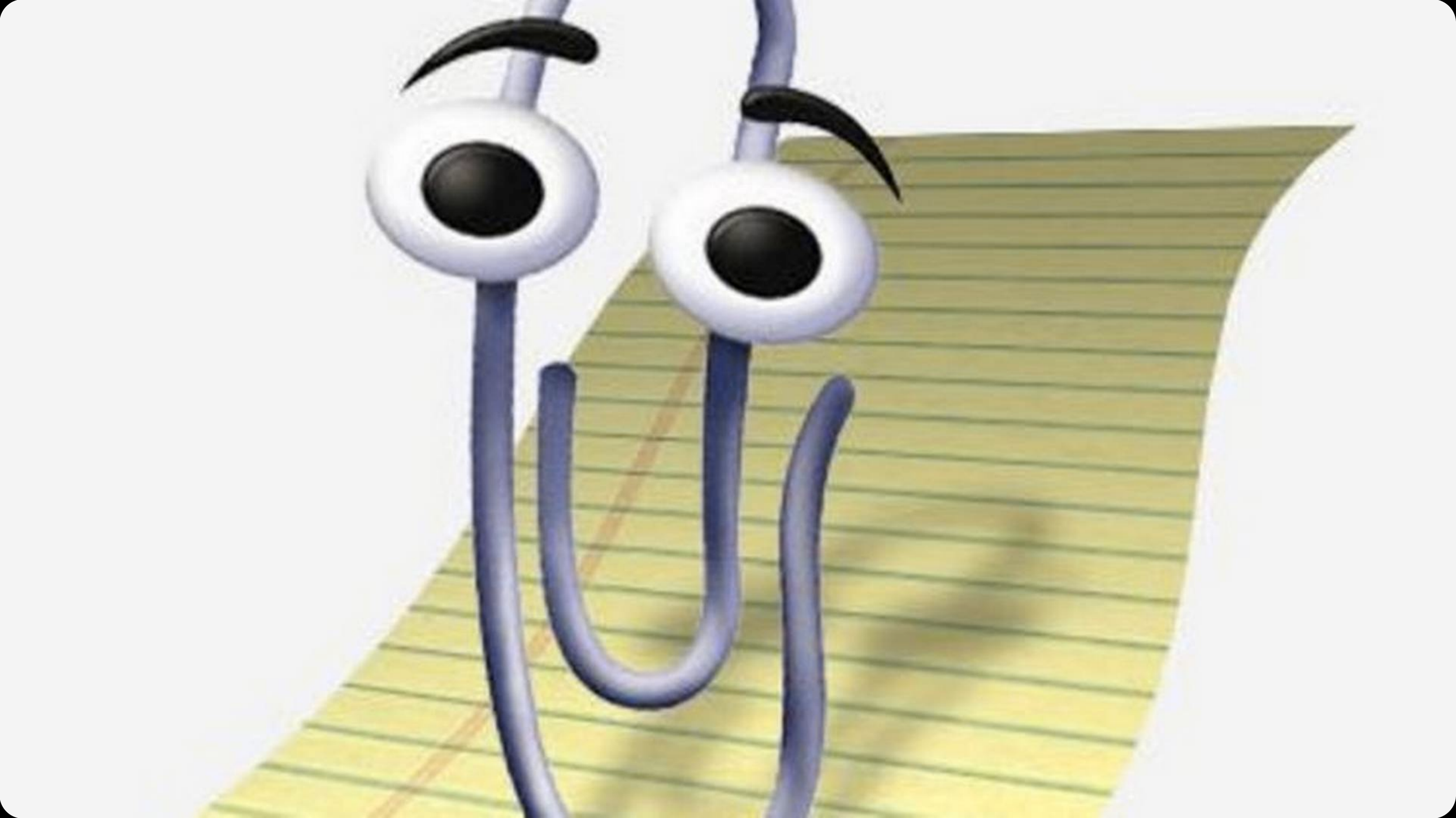
Protovis

vega

Data-Driven Documents

My software doesn’t know what I’m trying to do.

What if it did?



Demo document

For years I have been driving an old used car with a lot of mileage and I hate it. It gets me where I need to go, but I'm tired of fixing leaks and broken parts all the time. Its annoying every times I need to take it to the mechanic. Even when they take care of everything, I know I'll just end up going back there in a few weeks.

I have finally decided that I am not going to do it anymore. I have decided to buy a new car! Unfortunately, I have a problem. I have no idea what car to get. Do I want something fast? Do I want something big? Do I want something stylish? Something economical? I have so many choices that I don't even know where to begin. I am not sure if I will be able to make the decision on my own. I don't have not a lot of money, either, so I probably don't have many options.

After I did some research, I knew that I would need some expert advice. Eventually, I went to a local dealership to check out some new models. I

years,

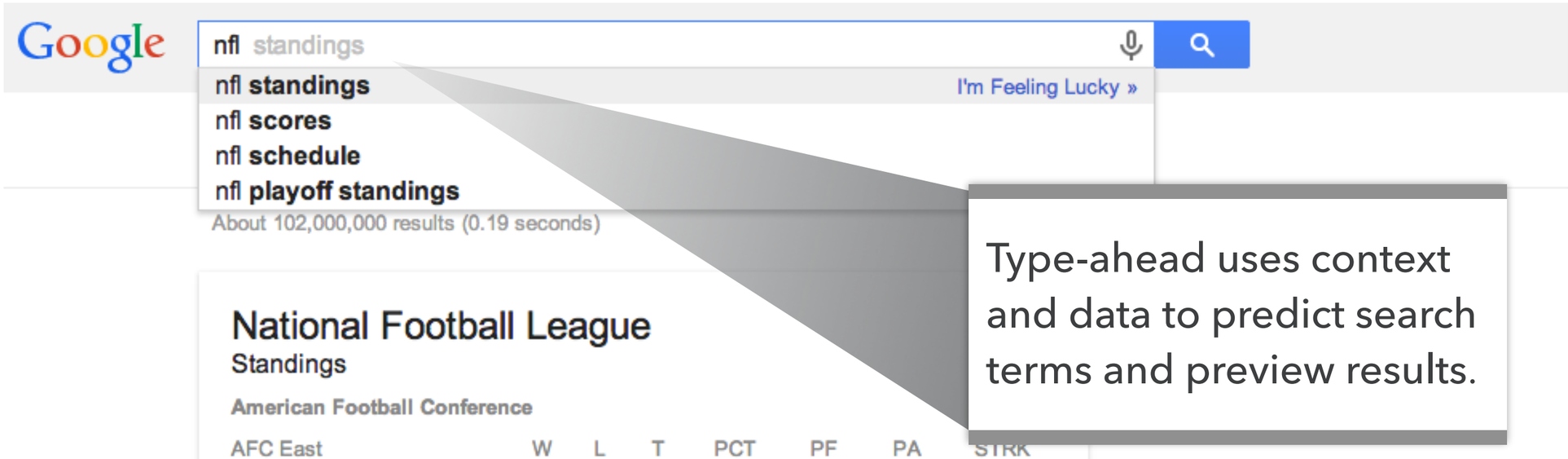
Possibly confused word: *Its*

every ~~times~~ → every time

~~not~~

~~did~~ → had done

Hints of Intelligent Interaction



National Football League

Standings

American Football Conference

AFC East	W	L	T	PCT	PF	PA	STRK
Patriots	12	4	0	.750	444	338	W2
Jets	8	8	0	.500	290	387	W2
Dolphins	8	8	0	.500	317	335	L2
Bills	6	10	0	.375	339	388	L1

News for **nfl standings**

NFL Power Rankings: Updated **Standings** Heading Toward 2014 Super Bowl

Bleacher Report - by David Daniels - 2 days ago

In one season, it digressed from having a Super Bowl-winning head coach and the **NFL's** most exciting player at QB to firing the coach

husky union building - Google

Secure | https://www.google.com/search?ei=clBNWvT5Fqvd0gKhiqX4Dg&q=husky+union+building&oq=husky+union&gs_l=psy-ab.3.0.0l10.6185.58769.0.61159.15.13.2.0.0.0.96.742.13.1...

Google

husky union building

husky union building **parking**
husky union building **hours**
4001 e stevens way ne, seattle, wa 98195
uw hub map

Report inappropriate predictions

The HUB - UW Departments Web Server - University of Washington

depts.washington.edu/thehub/

The **Husky Union Building** is one of several units within the Division of Student Life, is funded in part by the Services and Activities Fee (SAFC), and is comprised of 12 individual units including the Student Activities Office, **HUB** Games, **HUB** Event & Information Services, and the Resource Center among others. The **HUB** is ...

Directions
The HUB is located on upper campus. Allen Library is to the ...

About the HUB
The HUB supports the Husky Experience by Enhancing UW ...

Husky Den Food Court
Each Husky Den food area has different service hours during ...

HUB Hours
HUB Building Hours. Winter Quarter 2018: Jan 3 – Mar 16 ...

More results from washington.edu »

Husky Den - UW HFS - University of Washington

https://www.hfs.washington.edu/huskyden/

Husky Den. **Husky Union Building** (HUB) Phone: 206-616-5270. Centrally located in the **Husky Union Building** (HUB), Husky Den is a popular breakfast and lunch destination on campus. The food court is home to eight restaurants, a market and a variety of seating venues.

Campus Maps

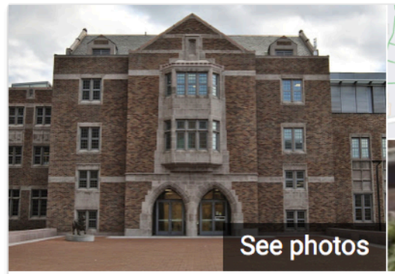
https://www.washington.edu/maps/

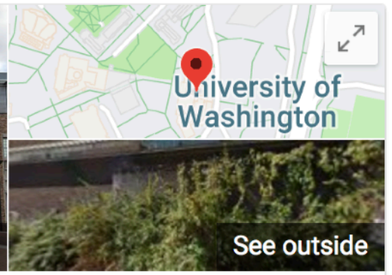
Explore. Computer Labs Food Gatehouses Landmarks Libraries Visitors Center Emergency Phones Parking Lots Photos Helpful links. Health Sciences Exp. South Lake **Union** ...

Husky Union Building - Wikipedia

https://en.wikipedia.org/wiki/Husky_Union_Building

Husky Union Building (The HUB) is a building at the University of Washington. It was opened in October 1949, and transferred from the Associated Students of the University of Washington (ASUW) to the university administration in April 1962. Construction began in July 2010 on a remodeling of the HUB. The Grand ...





Husky Union Building

4.5 ★★★★½ 111 Google reviews

Student union in Seattle, Washington

Website | Directions

Husky Union Building is a building at the University of Washington. It was opened in October 1949, and transferred from the Associated Students of the University of Washington to the university administration in April 1962. Wikipedia

Located in: University of Washington

Address: 4001 E Stevens Way NE, Seattle, WA 98195

Hours: Open today · 7AM–5:30PM

Phone: (206) 543-8191

Suggest an edit

Know this place? Answer quick questions

Questions & answers

Be the first to ask a question

Ask a question

alvin cheung - Google Search

Secure | https://www.google.com/search?ei=VVFNWt2hFOKa0wKeqIvQDQ&q=alvin+cheung&oq=alvin+cheung&gs_l=psy-ab.3..0l10.14495.28424.0.29216.13.13.0.0.0.0.112.778.11j2.13.0.....

Google

alvin cheung

- alvin cheung **uw**
- alvin cheung **dblp**
- alvin cheung **northstate**
- alvin cheung **nyu**
- alvin cheung **linkedin**
- alvin cheung **realtor**
- alvin cheung **razer**
- alvin cheung **new coast realty**
- alvin cheung **ropes**
- alvin cheung **facebook**

Report inappropriate predictions

Alvin Cheung | Computer Science & Engineering - UW CSE

https://www.cs.washington.edu/people/faculty/akcheung

I am an assistant professor in the Department of Computer Science & Engineering at the University of Washington, affiliated with the database and programming systems research groups. My research interests include program analysis, improving database application performance, and building big systems in general.

Alvin Cheung | UW College of Engineering

https://www.engr.washington.edu/facresearch/newfaculty/2014/alvincheung

Alvin Cheung will be joining UW CSE as a faculty member. He will receive his Ph.D. in Computer Science from MIT this fall. Alvin's research focuses on co-optimizing data-intensive applications by examining the database and the runtime system and environment together, which can enable order-of-magnitude speedups in ...

Alvin Cheung at University of Washington - RateMyProfessors.com

www.ratemyprofessors.com/ShowRatings.jsp?tid=2120358

Rating and reviews for Professor **Alvin Cheung** from University of Washington Seattle, WA United States.

Images for alvin cheung

Alvin Cheung

Author

Books: Computer-assisted Query Formulation

Feedback

dan suciu - Google Search

Secure | https://www.google.com/search?ei=sFBNWp_iGOmd0wL8gZSIAw&q=dan+suciu&oq=dan+suciu&gs_l=psy-ab.3..0l6j0i22i30k1l4.1901.163372.0.163786.11.11.0.0.0.0.95.599.11.11....

Google

dan suciu

- dan suciu **uw**
- dan suciu **dblp**
- dan suciu **rate my professor**
- dan suciu **bnr**
- dan suciu **cv**
- dan suciu **ubb**
- dan suciu **merrill lynch**
- dan suciu **money channel**
- dan suciu **linkedin**
- dan suciu **cass**

Report inappropriate predictions

Dan Suciu | Computer Science & Engineering - UW CSE - University ...
https://www.cs.washington.edu/people/faculty/suciu
I am a full professor in Computer Science & Engineering at the University of Washington. I teach databases and do research in data management. I apply formal theory to novel and difficult data management tasks. My past work has addressed various aspects of managing semistructured data, including query languages, ...

Dan Suciu | Computer Science & Engineering - UW CSE - University ...
https://www.cs.washington.edu/people/faculty/suciu/personal
There aren't that many Suciu's in Romania, and it is pretty uncommon in the USA: 91 Suciu's were listed at http://www.anywho.com, and 5 **Dan Suciu's**. (I ran these queries on anywho a few years ago. Now the site has "improved" and won't allow you to search for a person unless you specify a certain state. Welcome to the ...

Dan Suciu - Wikipedia
https://en.wikipedia.org/wiki/Dan_Suciu
Dan Suciu is a full professor of computer science at the University of Washington. He received his Ph.D. from the University of Pennsylvania in 1995 under the supervision of Val Tannen. After graduation, he was a principal member of the technical staff at AT&T Labs until he joined the University of Washington in 2000. Suciu ...

dblp: Dan Suciu
dblp.uni-trier.de › Persons
List of computer science publications by **Dan Suciu**.

Dan Suciu

Computer science researcher

Dan Suciu is a full professor of computer science at the University of Washington. He received his Ph.D. from the University of Pennsylvania in 1995 under the supervision of Val Tannen. Wikipedia

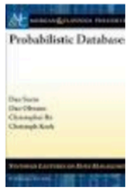
Education: University of Pennsylvania

Field: Computer Science

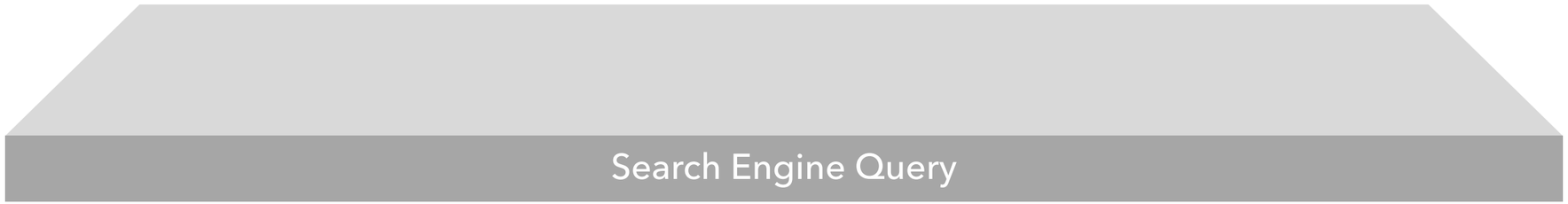
Awards: National Science Foundation CAREER Awards, Sloan Research Fellowship

Notable student: Mike Cafarella

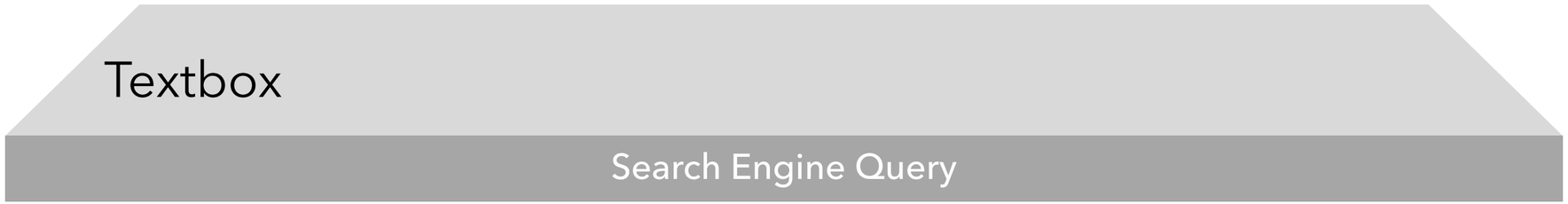
Books



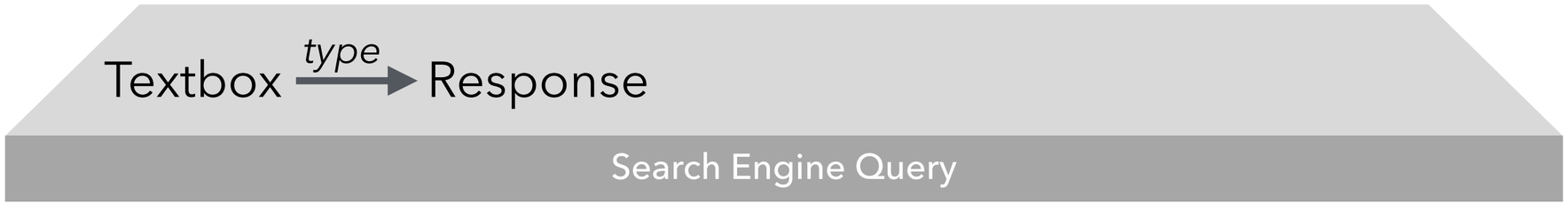
Search Query Auto-Complete



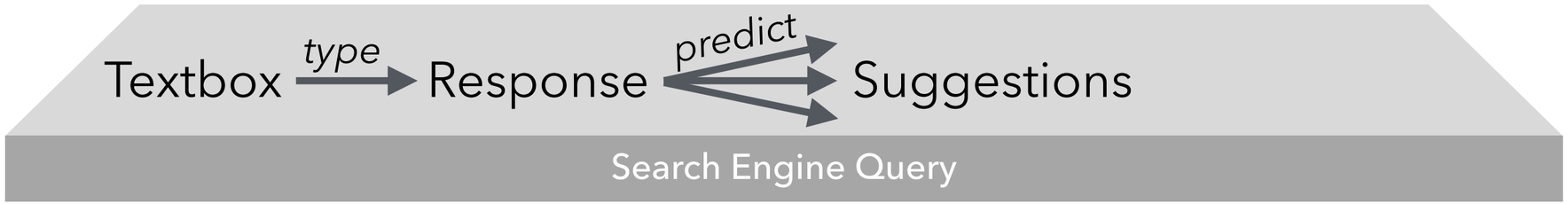
Search Query Auto-Complete



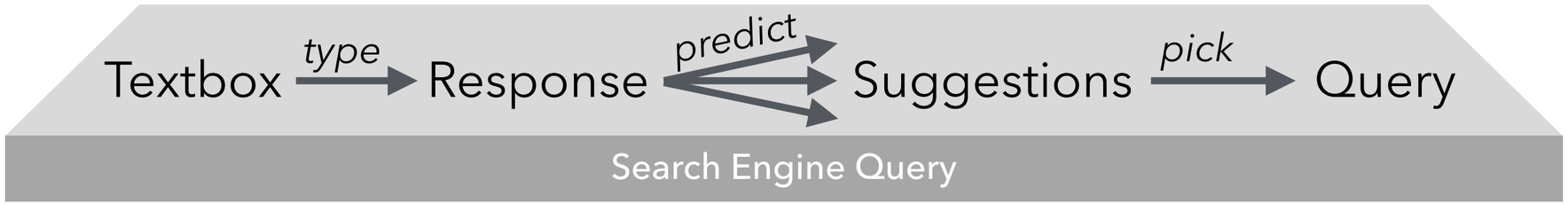
Search Query Auto-Complete



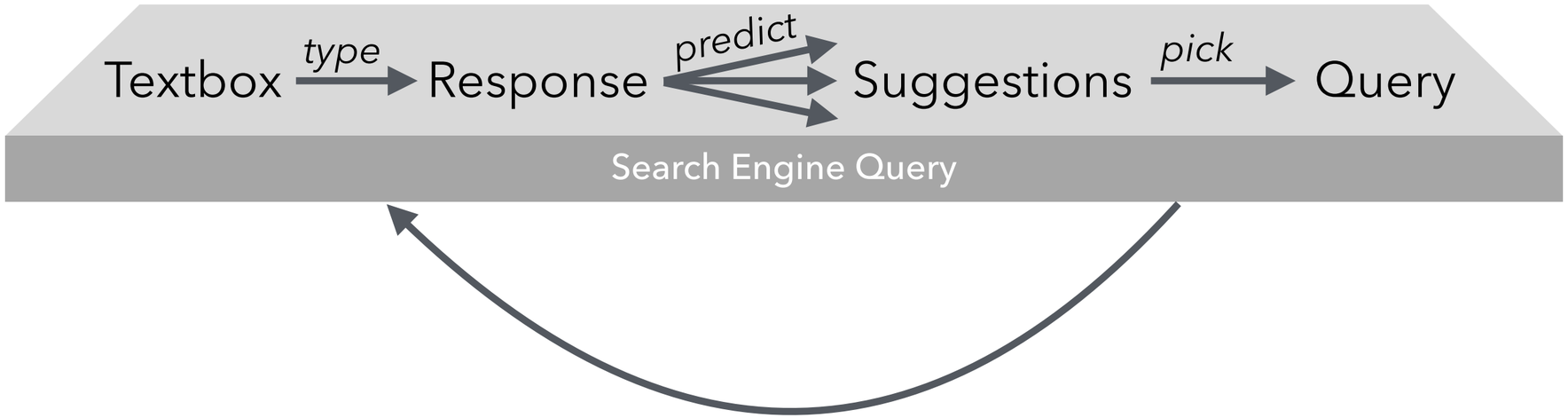
Search Query Auto-Complete



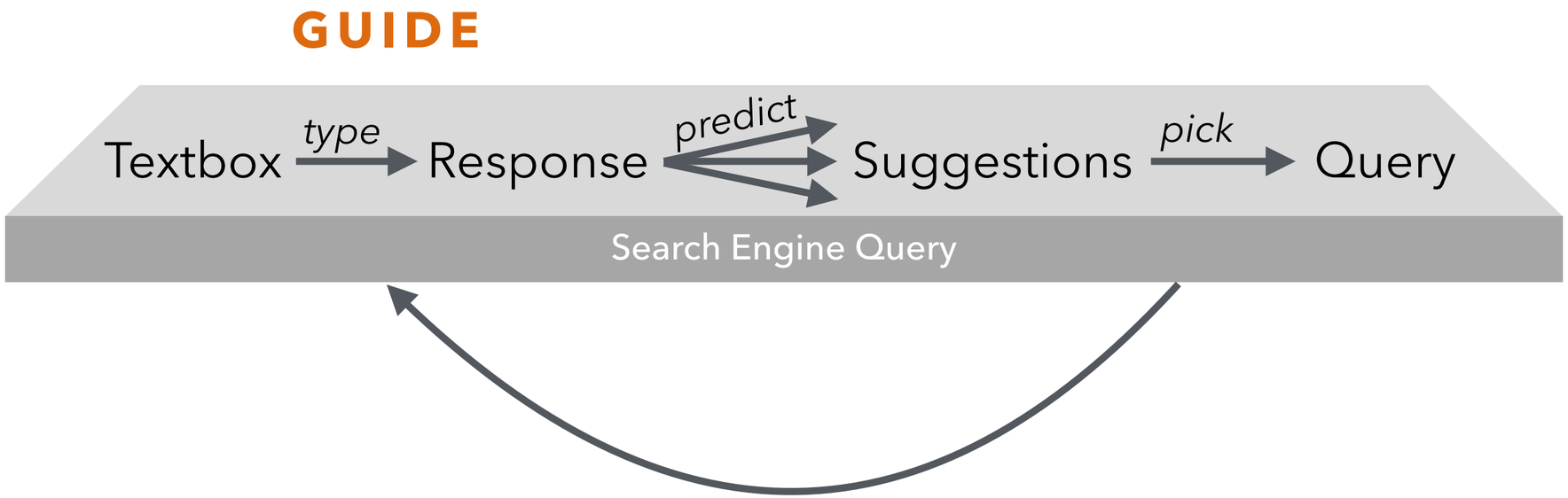
Search Query Auto-Complete



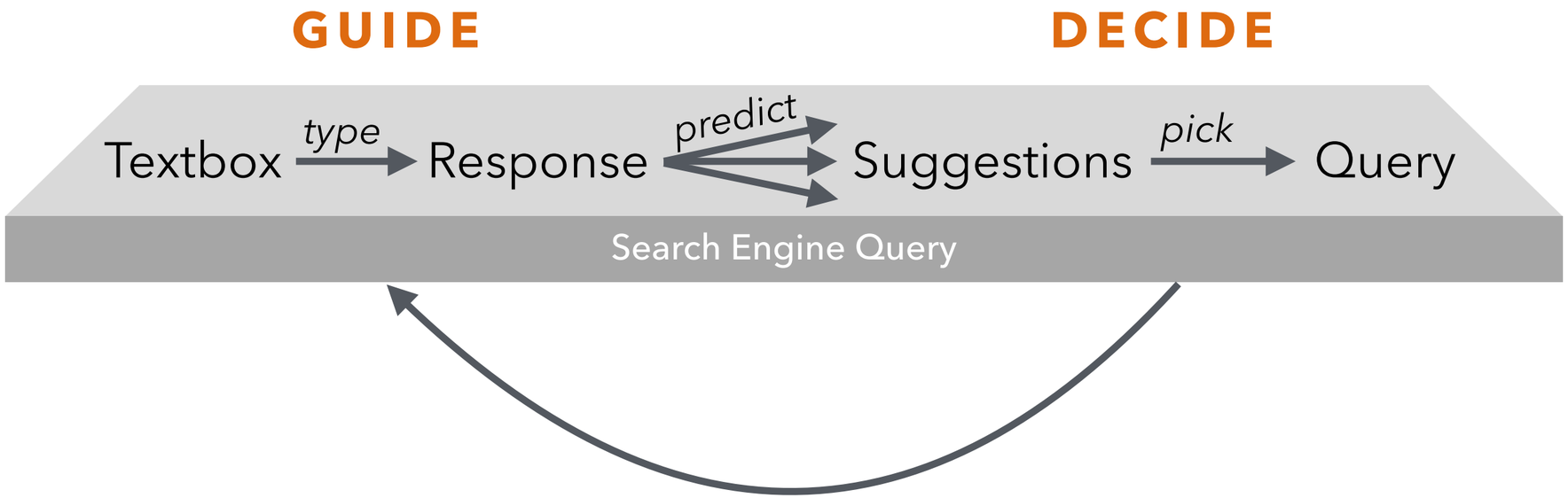
Search Query Auto-Complete



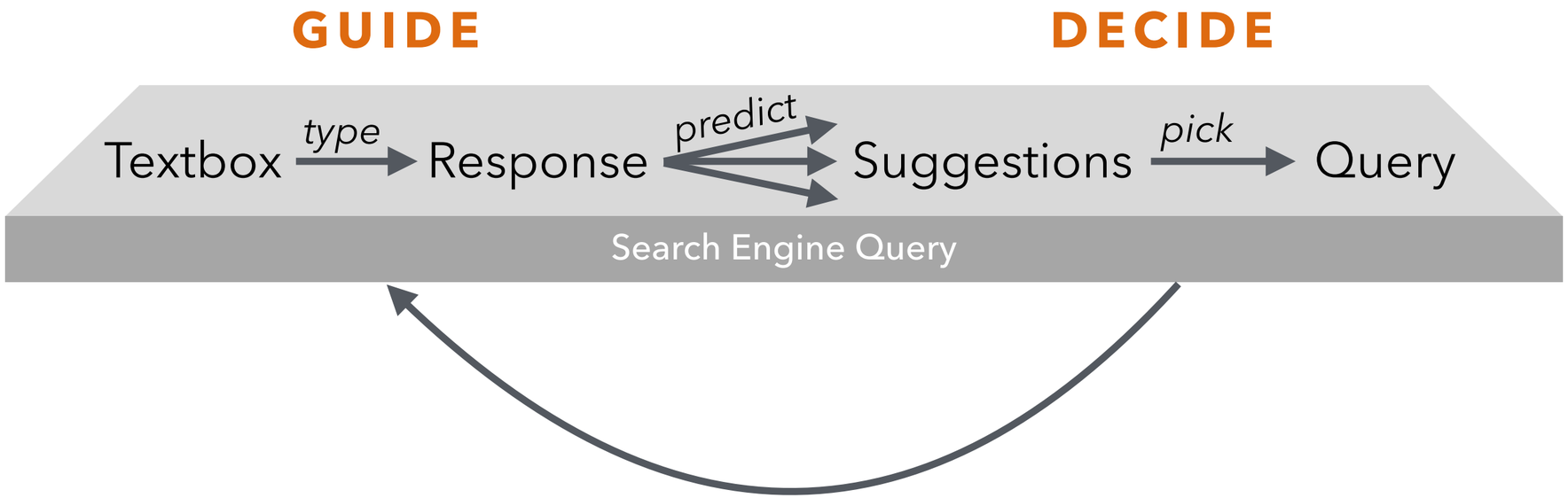
Search Query Auto-Complete



Search Query Auto-Complete

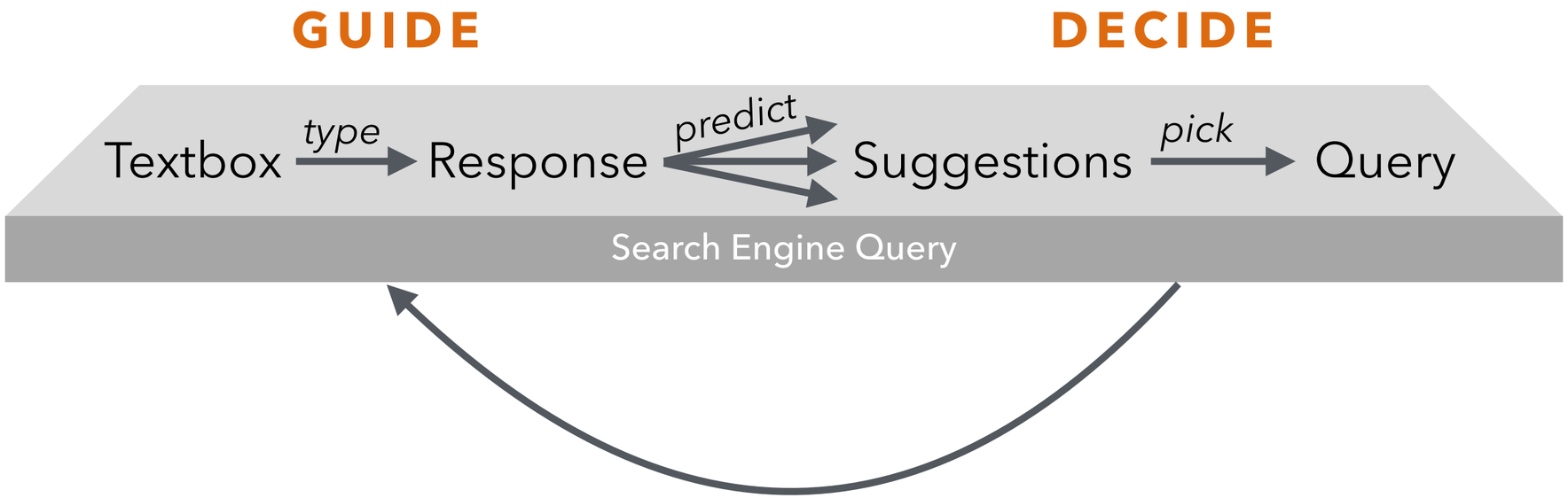


Search Query Auto-Complete



The input and output domains are the same: **text.**

Search Query Auto-Complete



What about more complex input/output relations?

Predictive Interaction Goals

Accelerate successful task completion.

Scale to large data or batch repetition.

Support discovery and ambiguous intent.

Long-term learning and improvement.

Predictive Interaction Goals

Accelerate successful task completion.

Scale to large data or batch repetition.

Support discovery and ambiguous intent.

Long-term learning and improvement.

Strategy

Model user interface actions in a **domain-specific language (DSL)**. Leverage the language to

(1) predict potential actions, and

(2) decouple UI from underlying runtime.

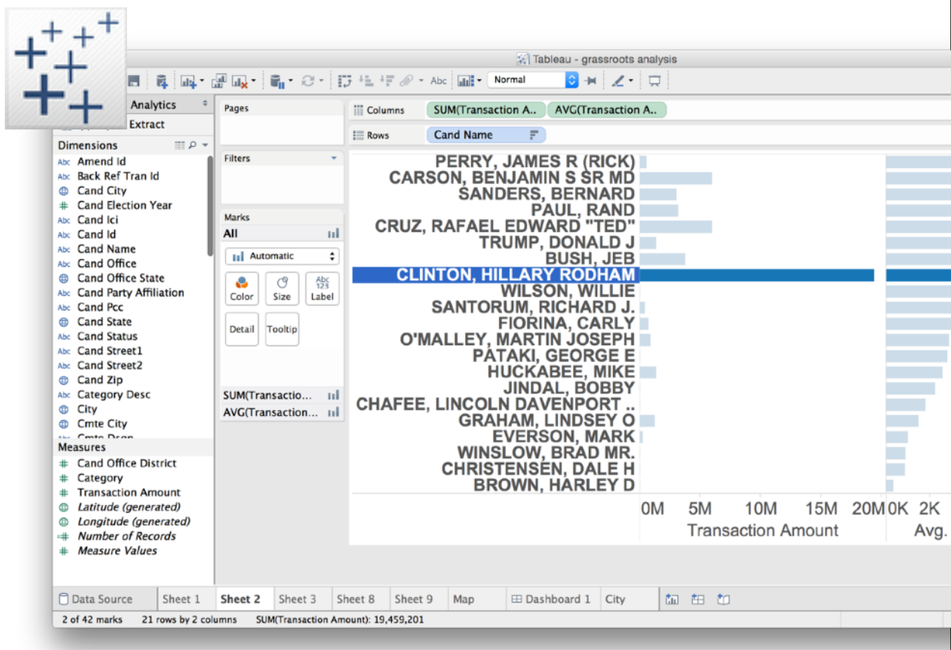
Predictive Interaction Goals

Accelerate successful task completion.

Scale to large data or batch repetition.

Support discovery and ambiguous intent.

Long-term learning and improvement.



Strategy

Model user interface actions in a **domain-specific language (DSL)**. Leverage the language to

(1) predict potential actions, and

(2) decouple UI from underlying runtime.

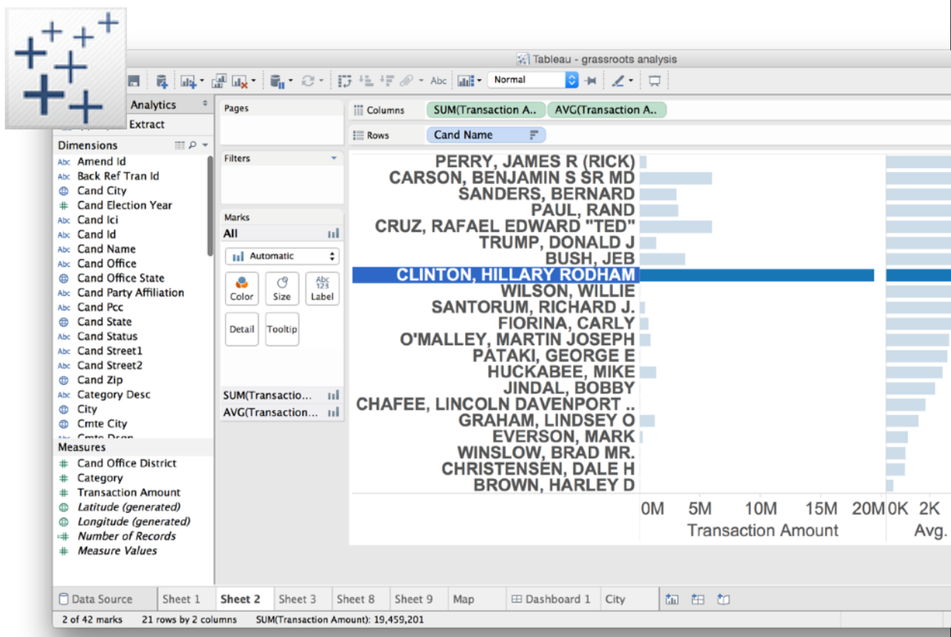
Predictive Interaction Goals

Accelerate successful task completion.

Scale to large data or batch repetition.

Support discovery and ambiguous intent.

Long-term learning and improvement.

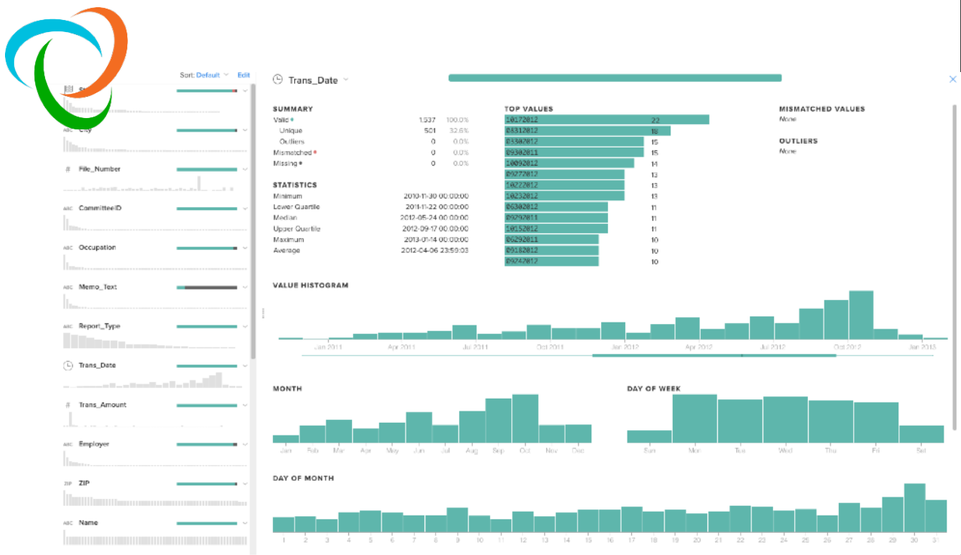


Strategy

Model user interface actions in a **domain-specific language (DSL)**. Leverage the language to

(1) predict potential actions, and

(2) decouple UI from underlying runtime.



EXAMPLE:
Data Visualization

Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category

- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region

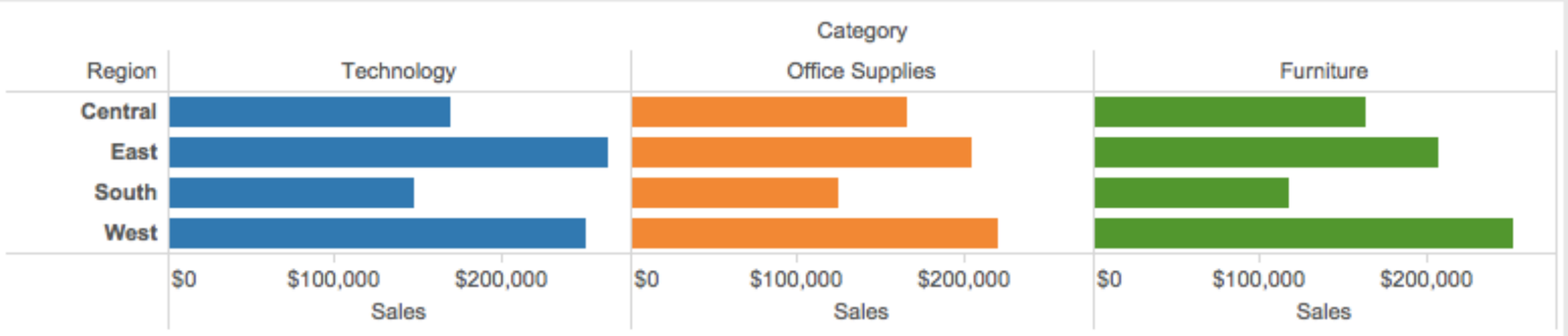


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category

- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region

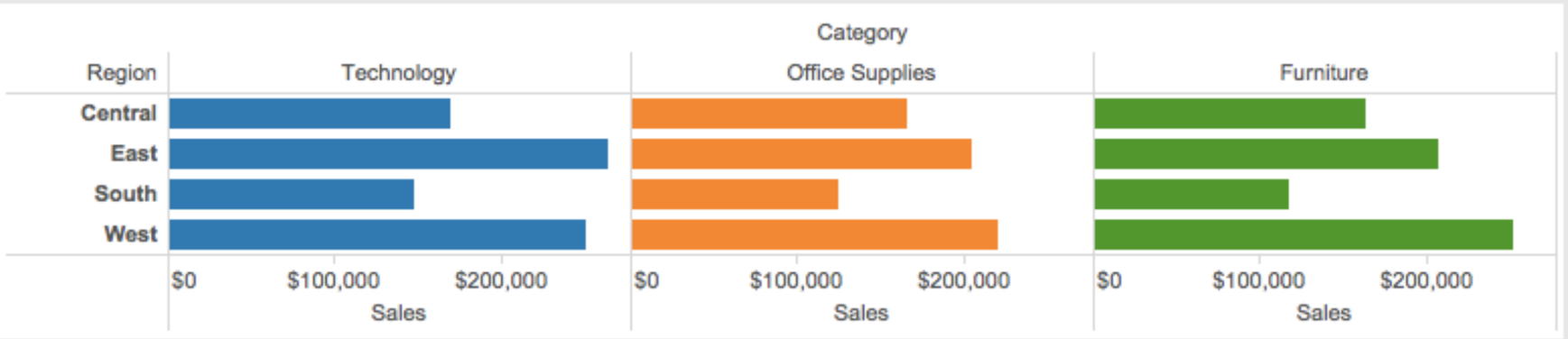


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category

- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region

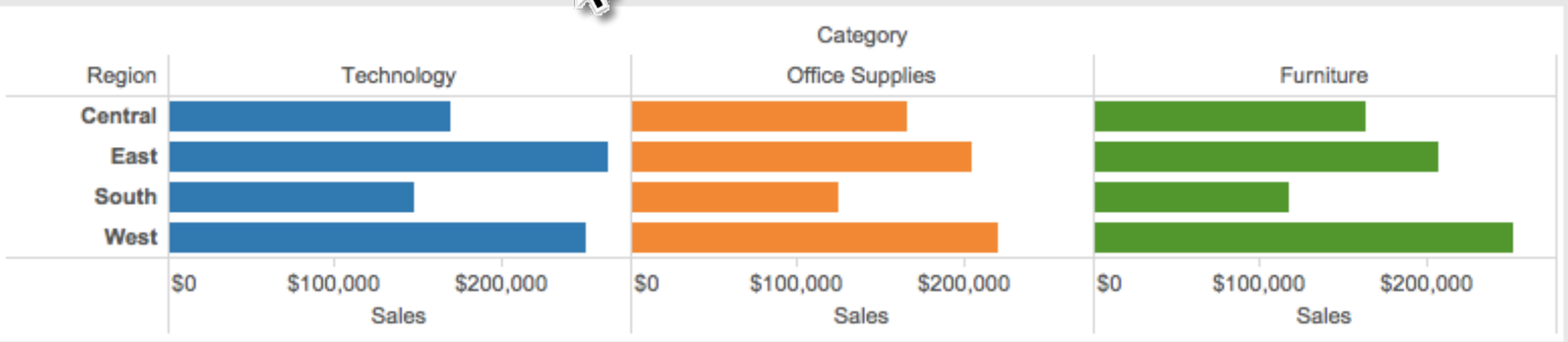


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category
- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region | Segment

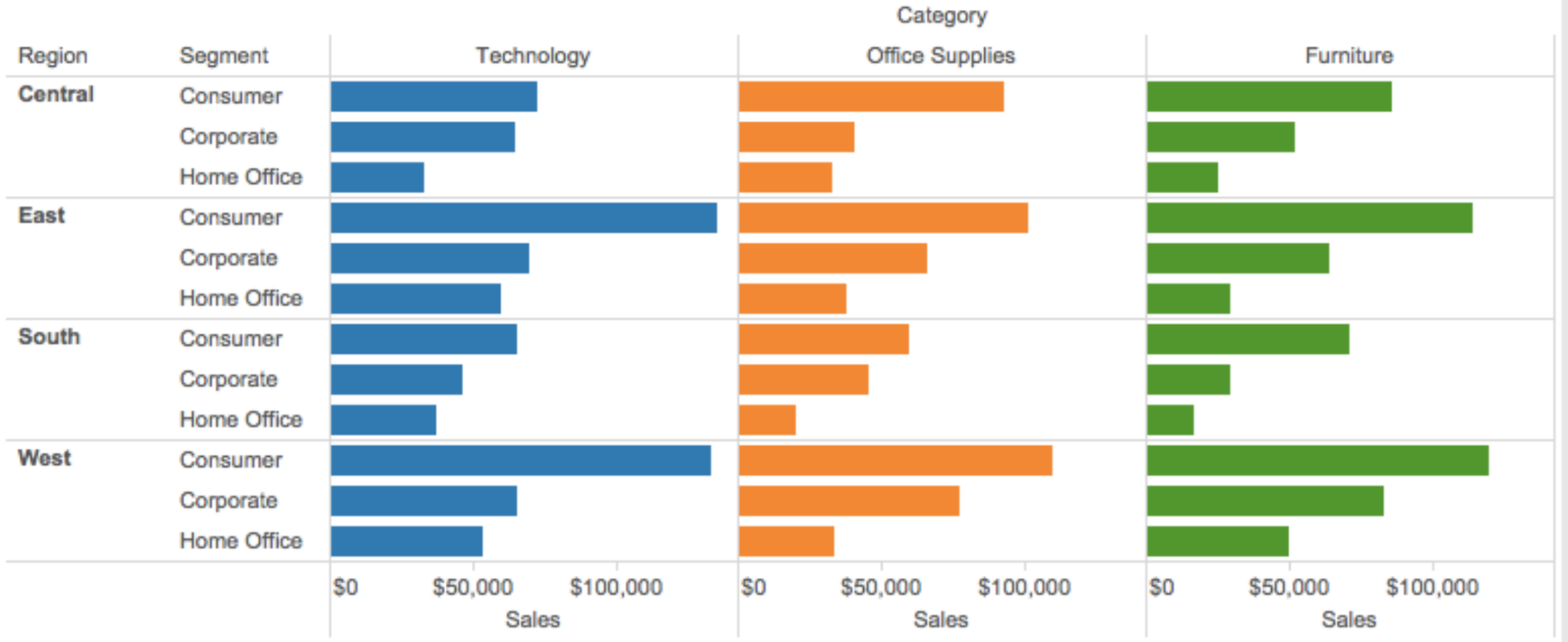


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category
- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region | Segment

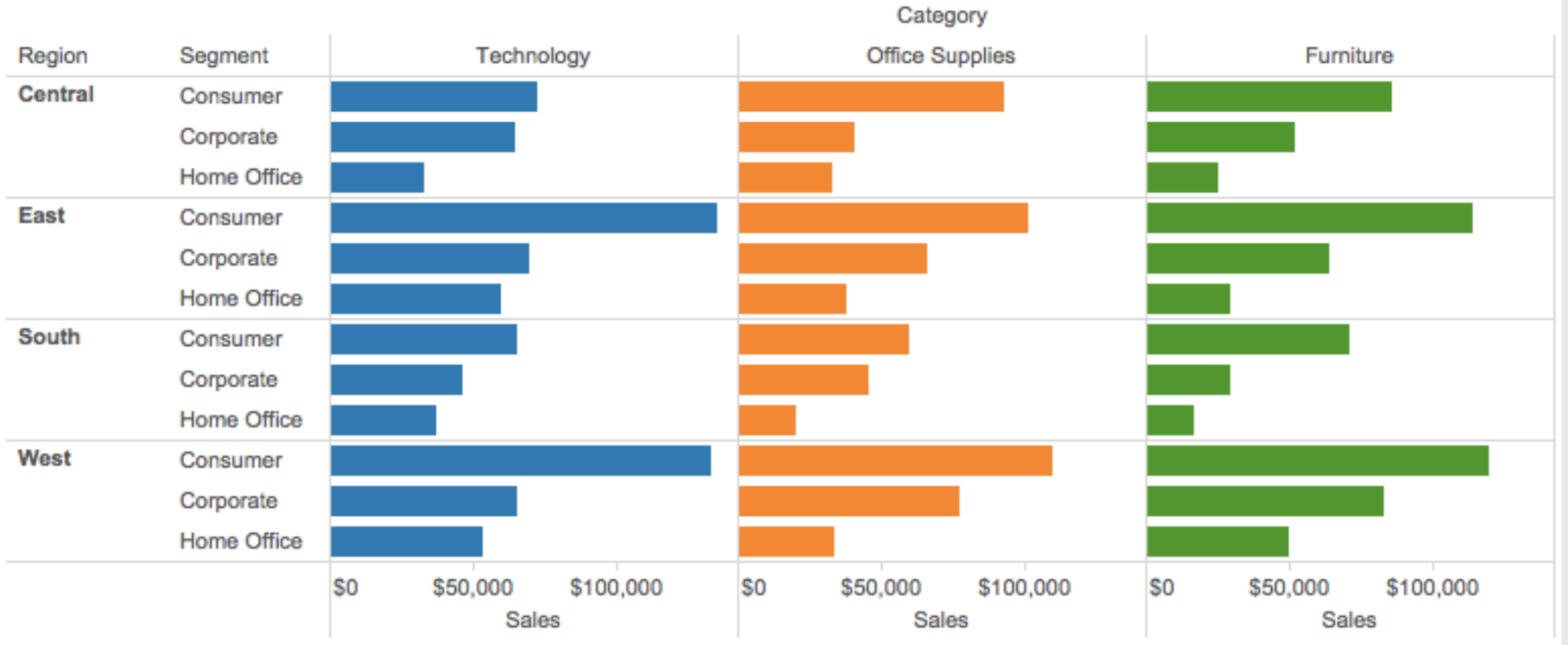


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

Category

Category
- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales)

Rows: Region | Segment

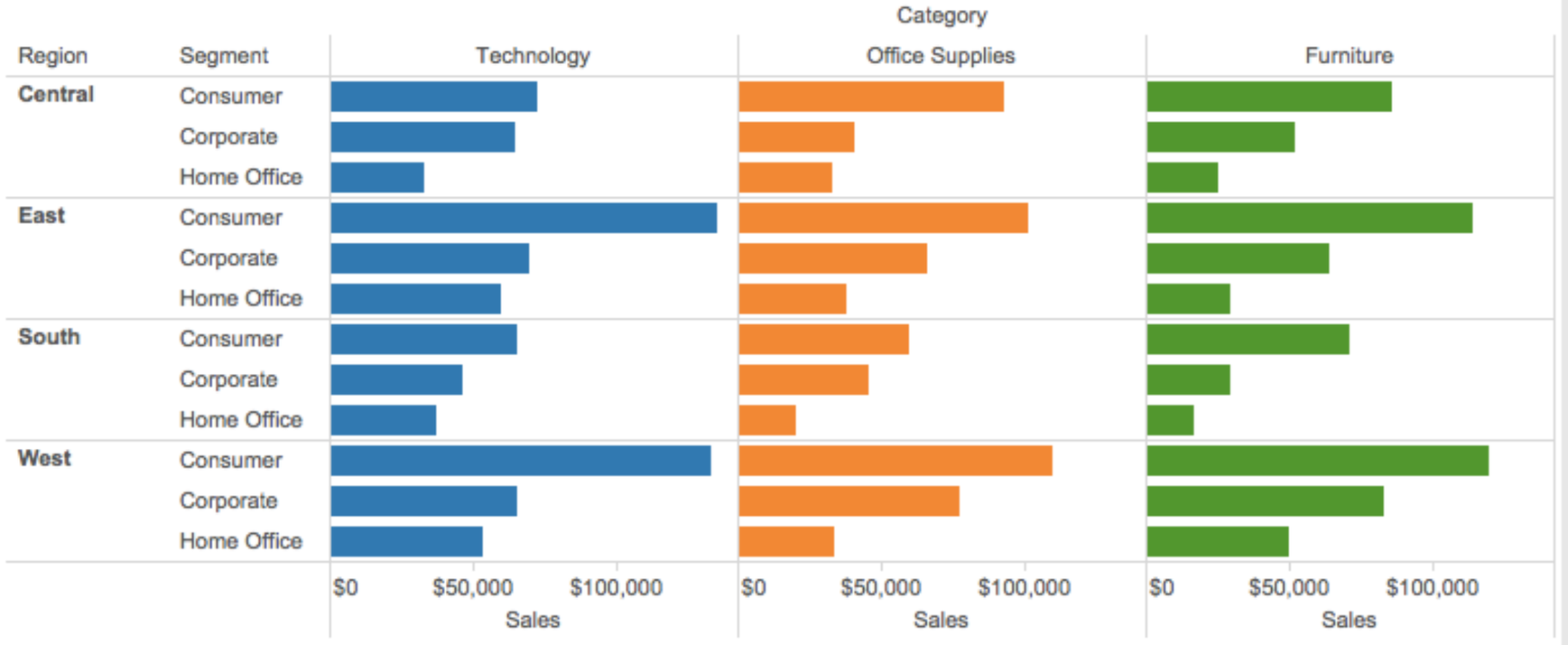


Tableau - Book1

Normal

Show Me

Data | Analytics

Sample – Superstore

Dimensions

- Customer
  - Customer Name
  - Segment
- Order
- Location
- Product
  - Category
  - Sub-Category
  - Manufacturer
  - Product Name
- Profit (bin)
- Region
- *Measure Names*

Measures

- Discount
- Profit
- Profit Ratio
- Quantity
- Sales
- *Latitude (generated)*
- *Longitude (generated)*
- *Number of Records*
- *Measure Values*

Pages

Filters

Marks

All

Automatic

Color | Size | Label | Detail | Tooltip

Category

SUM(Sales)

SUM(Profit)

Category

- Technology
- Office Supplies
- Furniture

Columns: Category | SUM(Sales) | SUM(Profit)

Rows: Region | Segment

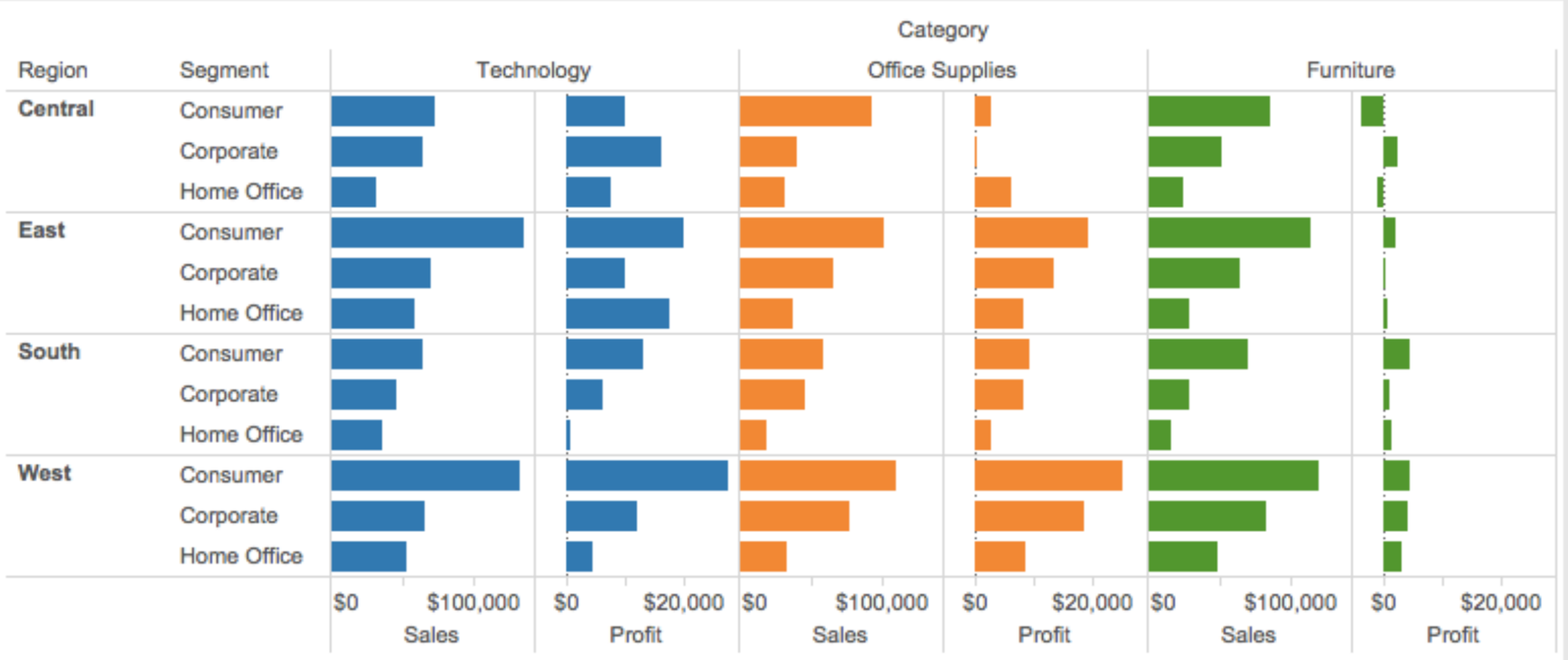


Tableau - Book1

Normal

Show Me

Columns: Category | SUM(Sales) | SUM(Profit)

Rows: Region | Segment

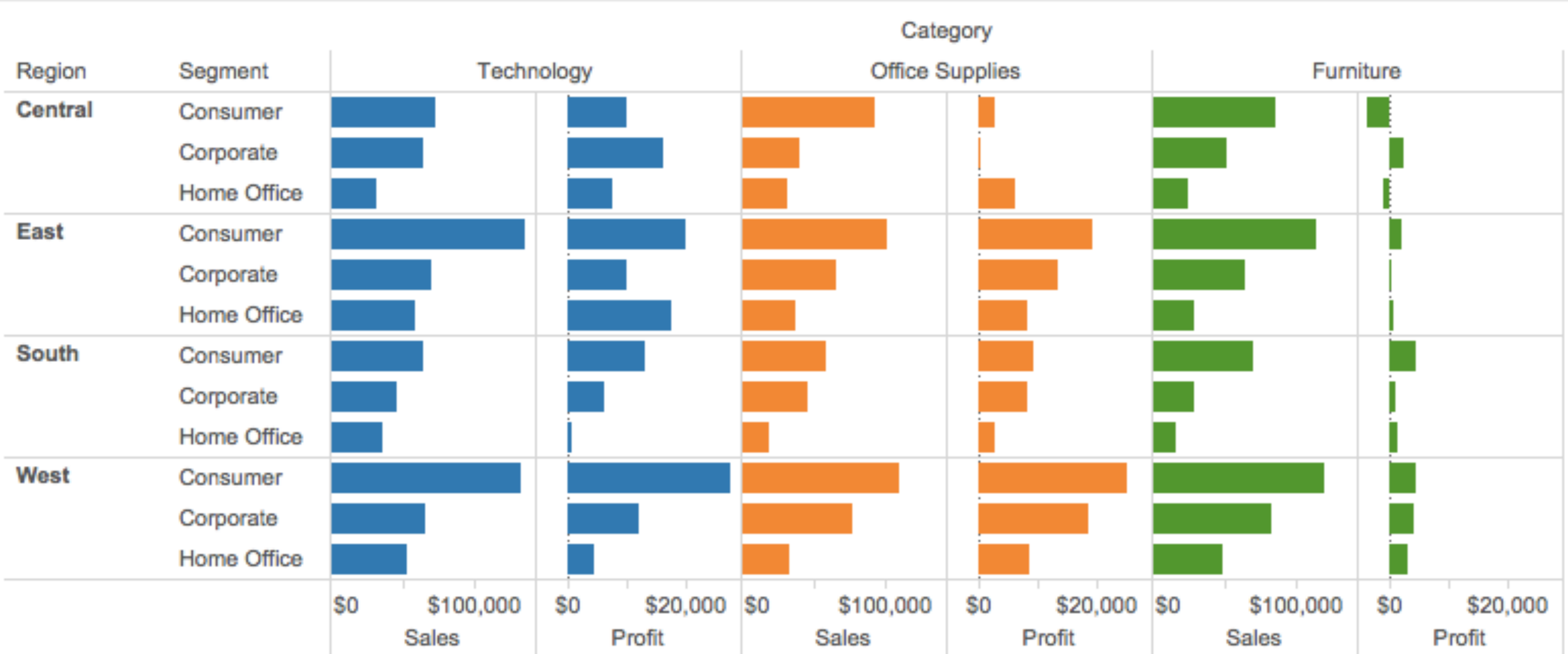


Tableau - Book1

Normal

Show Me

Columns: Category x SUM(Sales) + SUM(Profit)

Rows: Region \ Segment

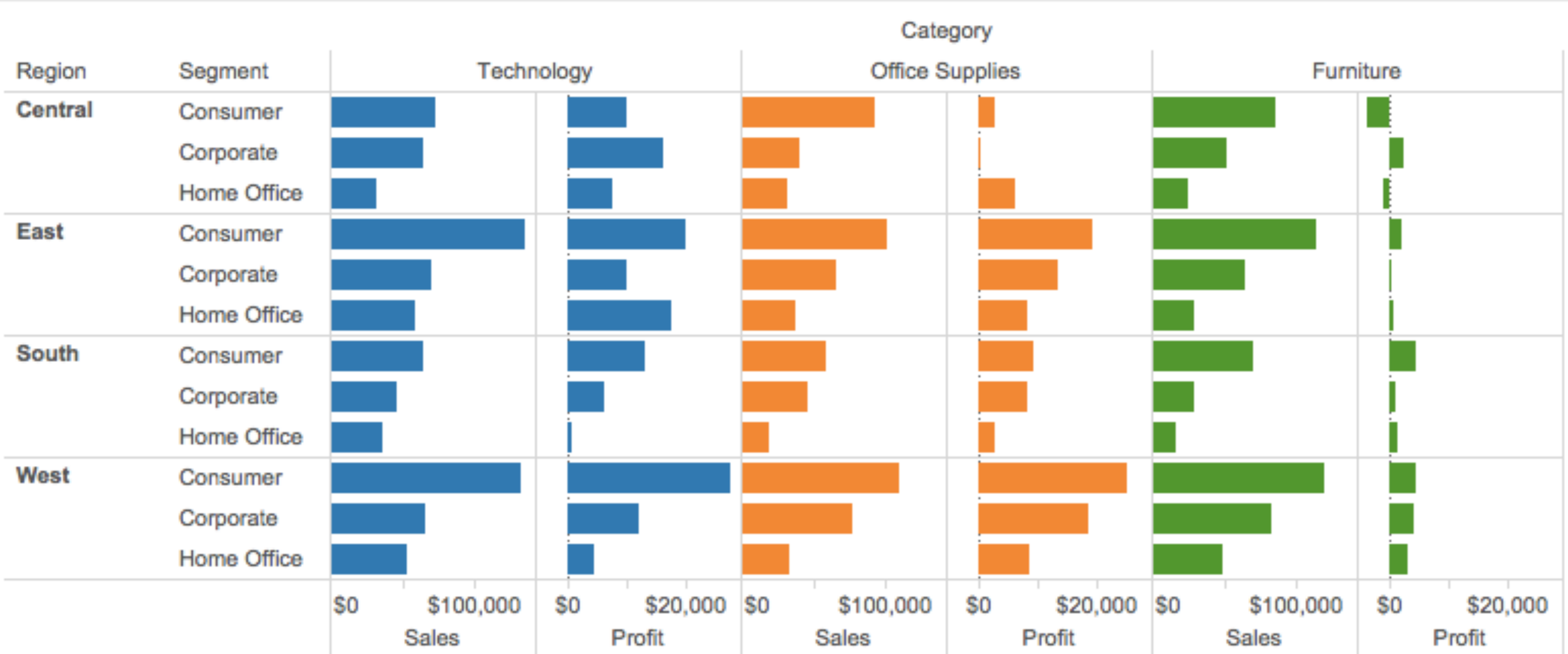


Tableau - Book1

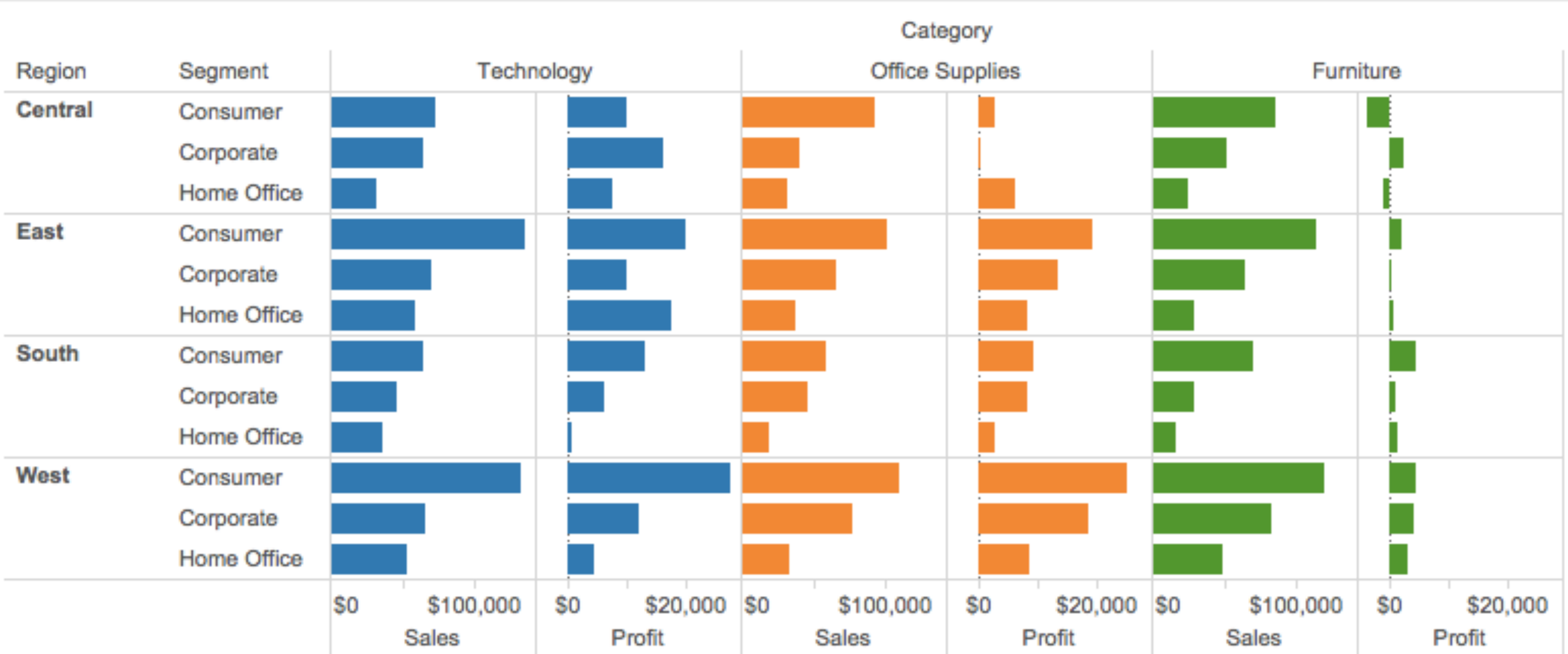
Normal

Show Me

Columns: Category X SUM(Sales) + SUM(Profit)

Rows: Region \ Segment

–> GROUP BY Category, Region, Segment



VizQL: A DSL for Tabular Visualization

VizQL: A DSL for Tabular Visualization

Operators:

concatenation (+)

cross product (x)

nest (\)

Operands:

Ordinal fields

Quantitative fields

VizQL: A DSL for Tabular Visualization

Operators:
concatenation (+)
cross product (x)
nest (\)

Operands:
Ordinal fields
Quantitative fields

The operators (+, x, \) and operands (O, Q) provide an **algebra for tabular visualization**.

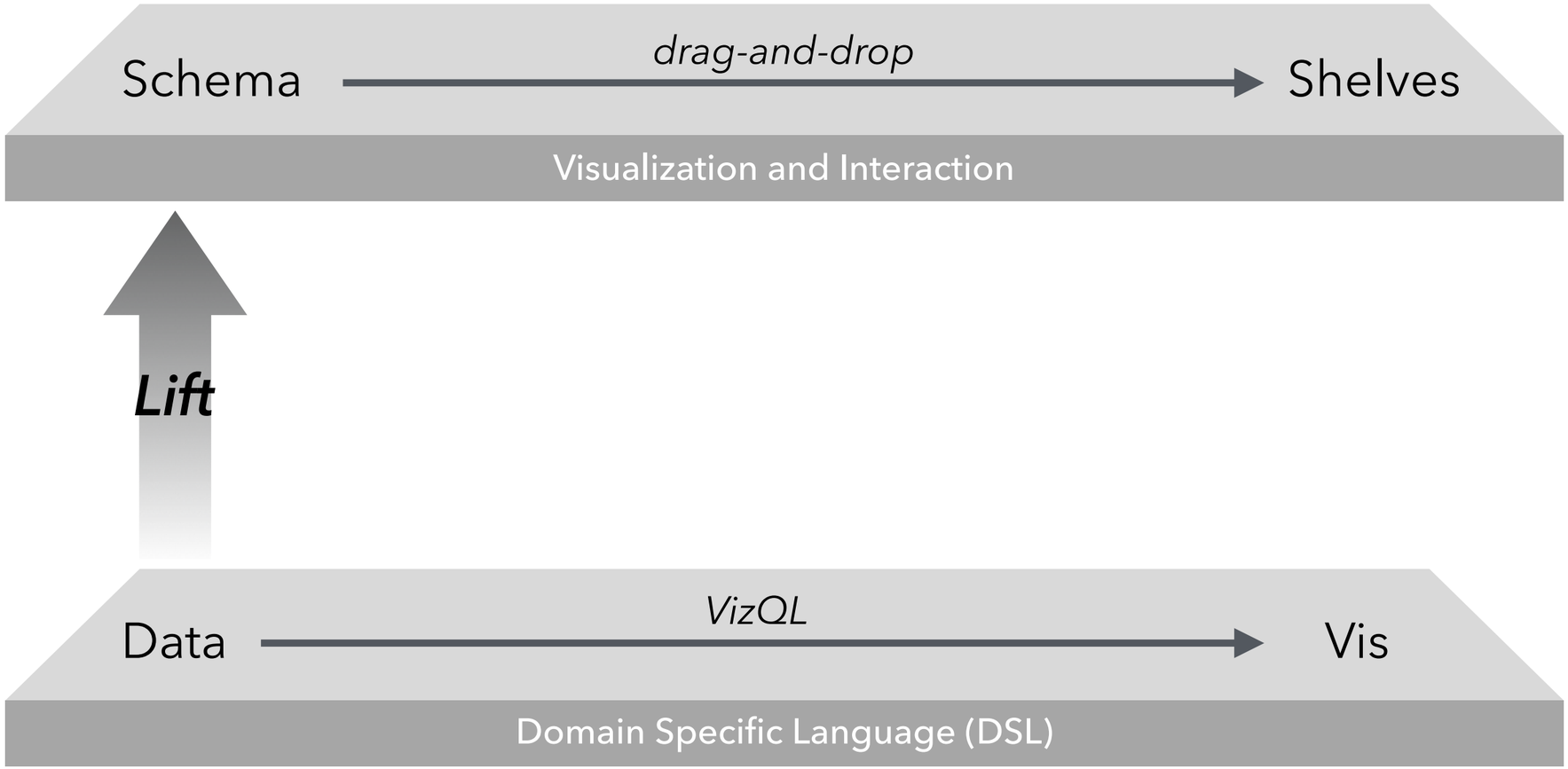
Algebraic statements are then **compiled** to:
Queries: selection, projection, group-by…
Visualizations: partitions, visual encodings

Users make statements via **drag-and-drop**
This specifies parameters, *not* operators!
Operators inferred by data type (O, Q)

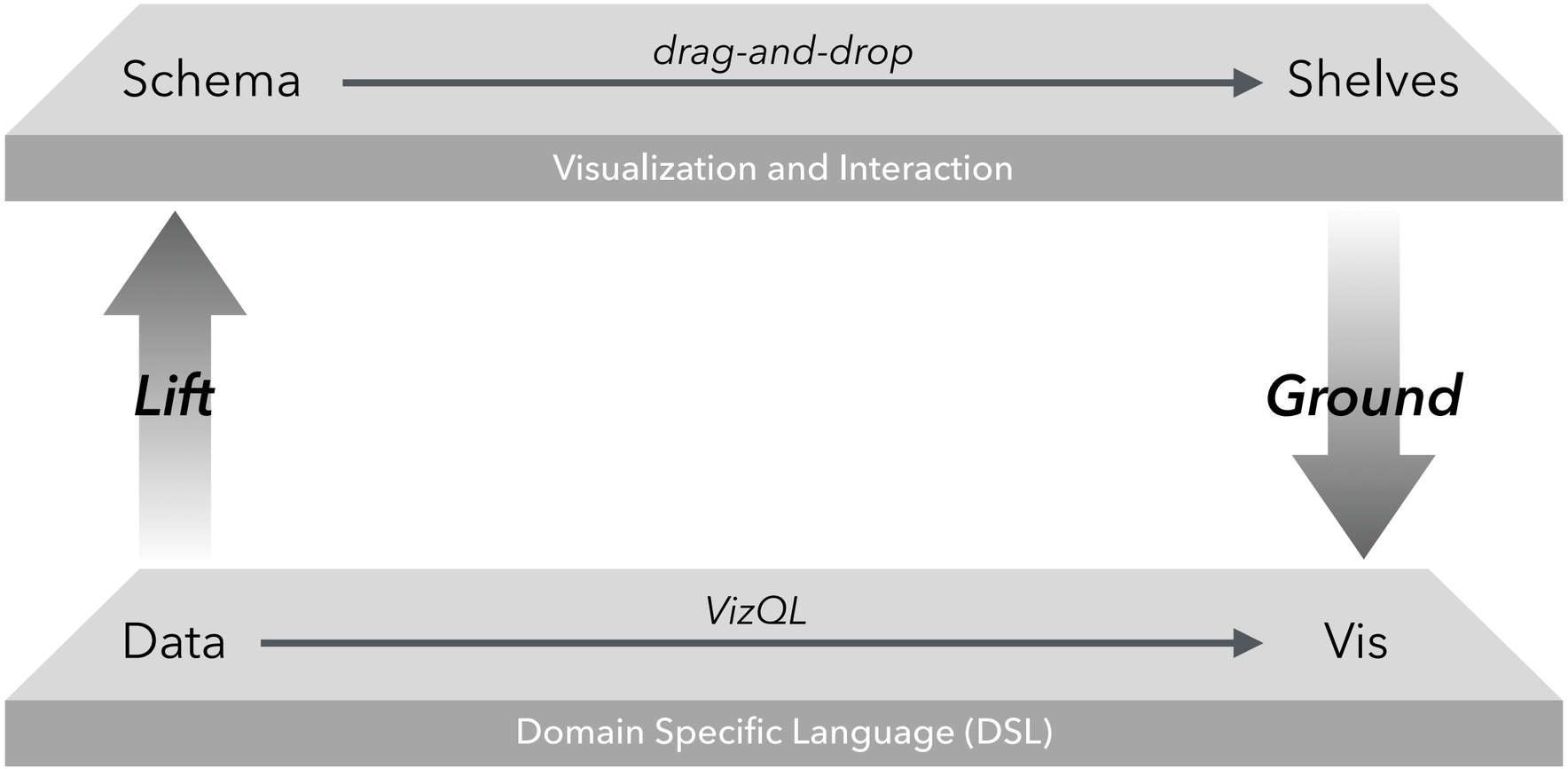
Mapping from Textual to Visual Languages

Mapping from Textual to Visual Languages

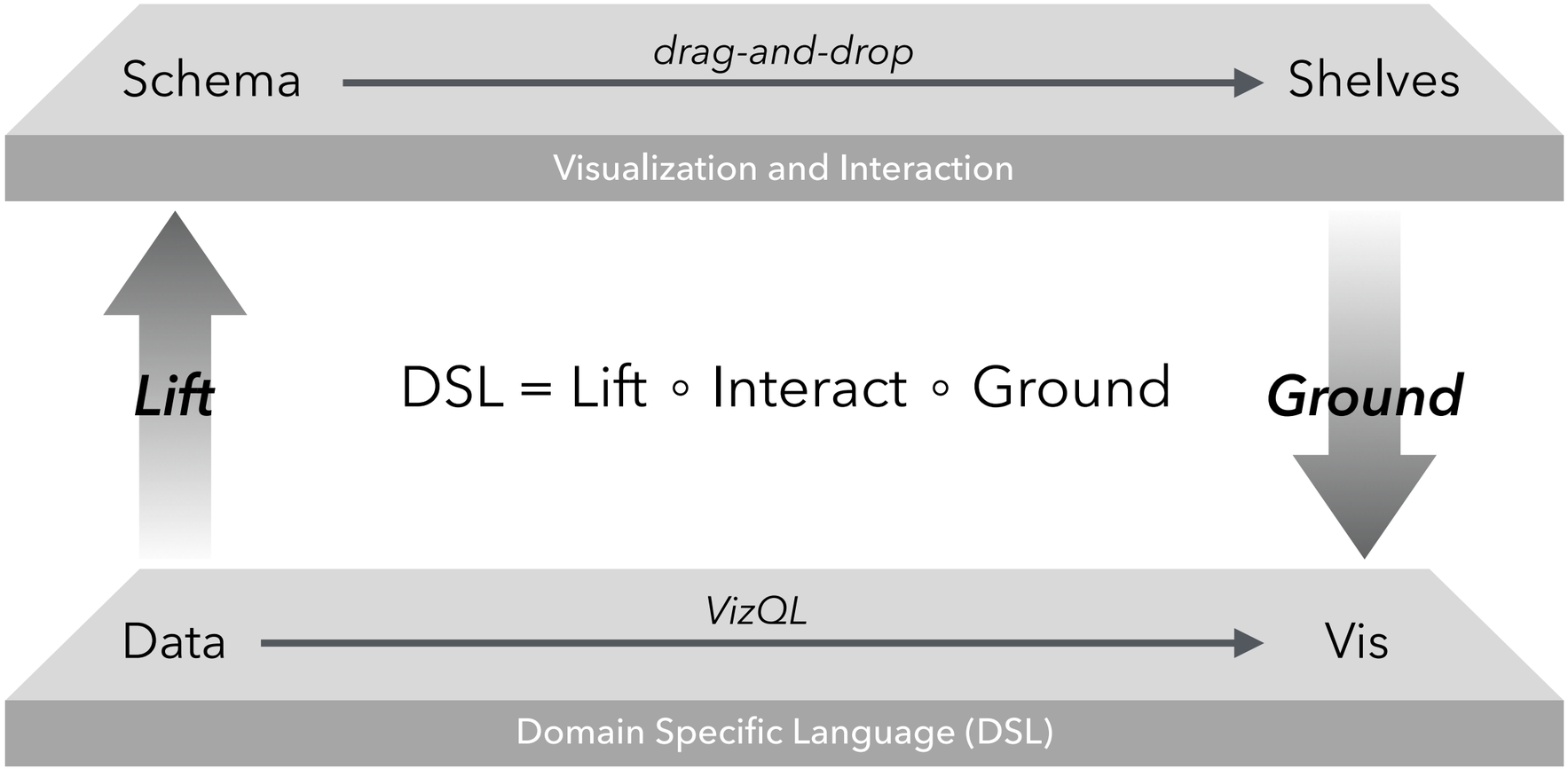
Mapping from Textual to Visual Languages



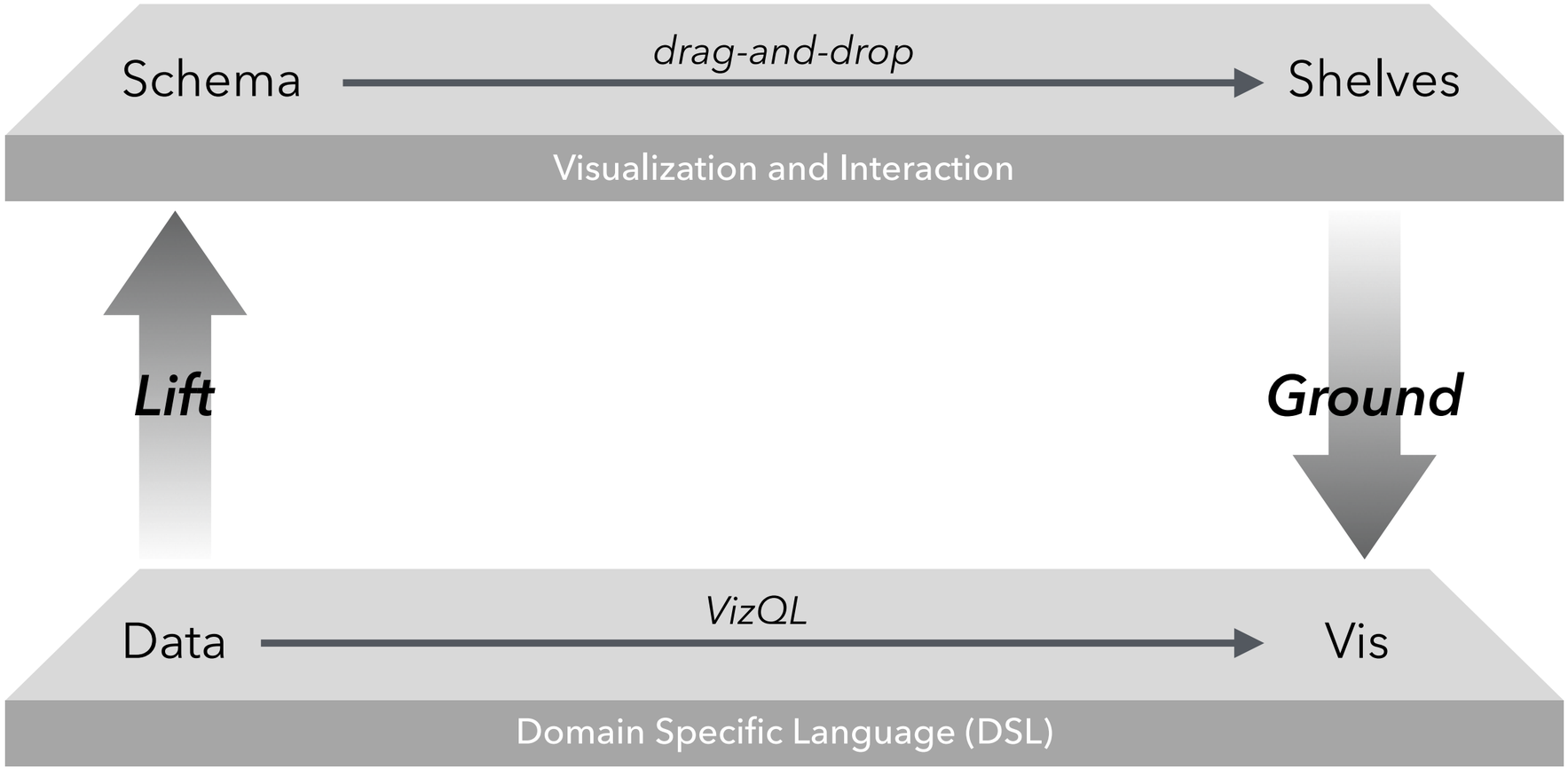
Mapping from Textual to Visual Languages



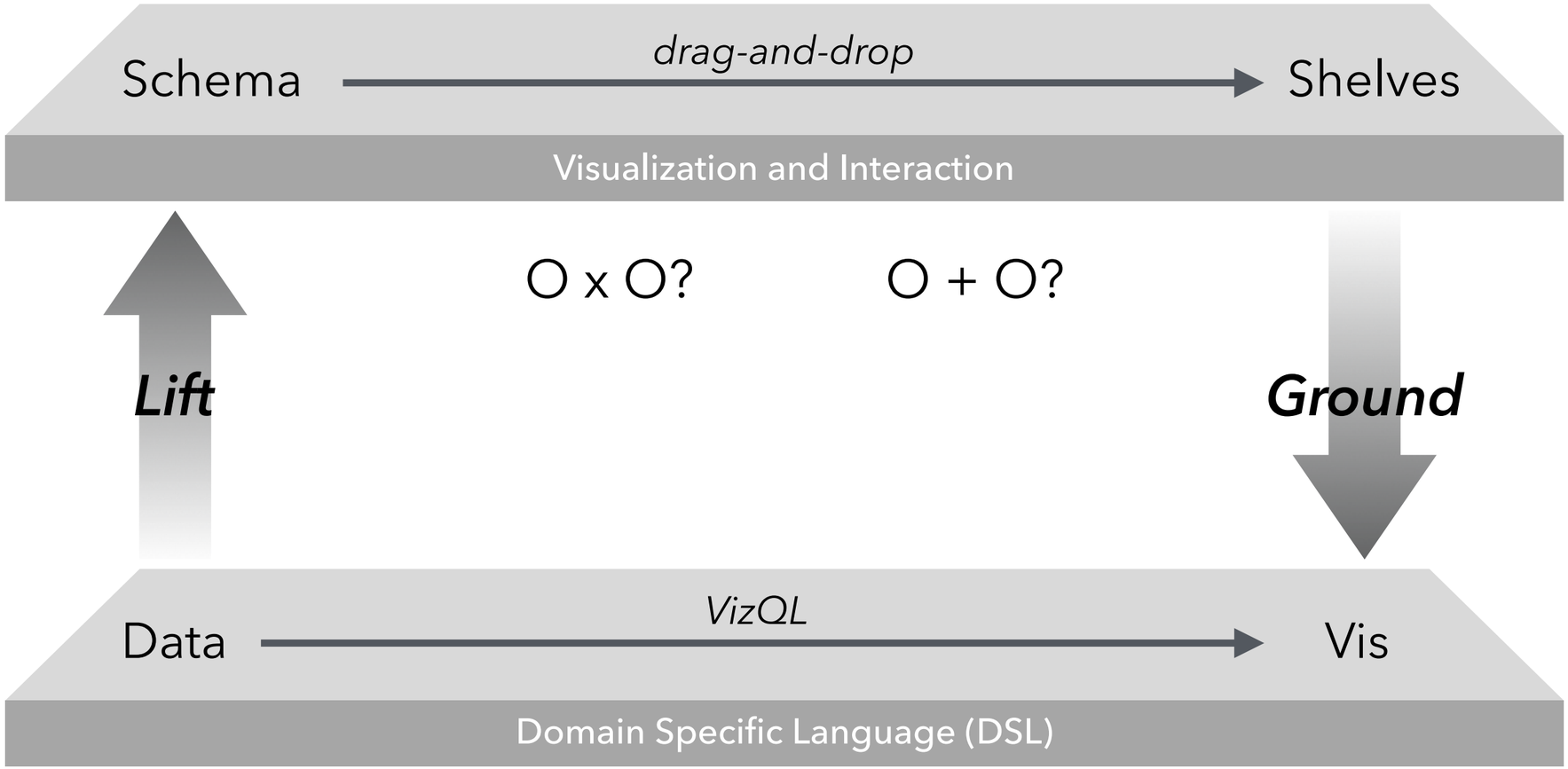
Mapping from Textual to Visual Languages



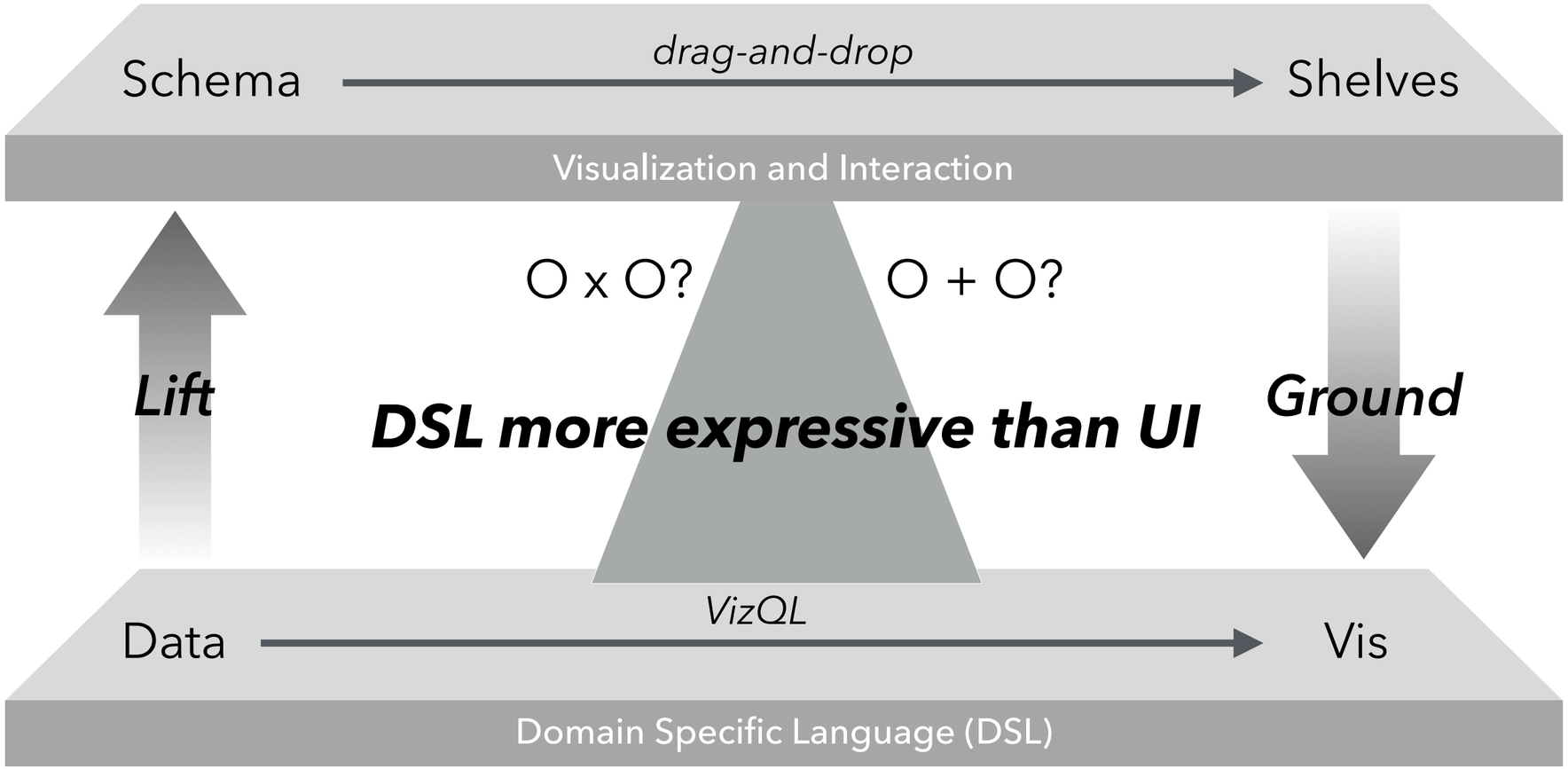
Are the Languages Isomorphic?



Are the Languages Isomorphic?



Are the Languages Isomorphic?



EXAMPLE:

Data Wrangling

I spend more than half of my time integrating, cleansing and transforming data without doing any actual analysis. Most of the time I'm lucky if I get to do any "analysis" at all.

Anonymous Data Scientist
from our 2012 interview study

Big Data Borat

@BigDataBorat

Following

In Data Science, 80% of time spent prepare data, 20% of time spent complain about need for prepare data.

Bureau of Justice Statistics - Data Online
http://bjs.ojp.usdoj.gov/

Reported crime in Alabama

Year	Population		Property crime rate		Burglary rate	Larceny-theft rate	Motor vehicle theft rate
2004	4525375	4029.3	987	2732.4	309.9		
2005	4548327	3900	955.8	2656	289		
2006	4599030	3937	968.9	2645.1	322.9		
2007	4627851	3974.9	980.2	2687	307.7		
2008	4661900	4081.9	1080.7	2712.6	288.6		

Reported crime in Alaska

Year	Population		Property crime rate		Burglary rate	Larceny-theft rate	Motor vehicle theft rate
2004	657755	3370.9	573.6	2456.7	340.6		
2005	663253	3615	622.8	2601	391		
2006	670053	3582	615.2	2588.5	378.3		
2007	683478	3373.9	538.9	2480	355.1		
2008	686293	2928.3	470.9	2219.9	237.5		

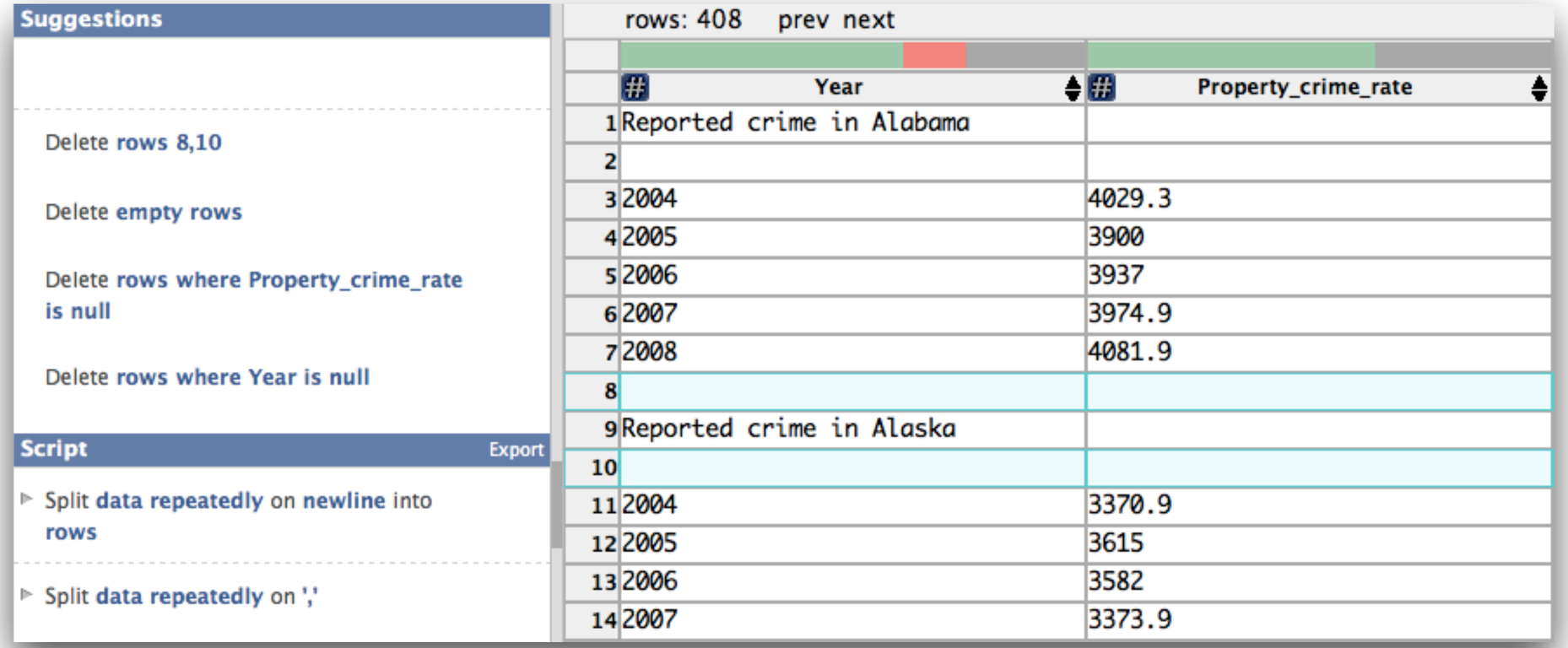
Reported crime in Arizona

Year	Population		Property crime rate		Burglary rate	Larceny-theft rate	Motor vehicle theft rate
2004	5739879	5073.3	991	3118.7	963.5		
2005	5953007	4827	946.2	2958	922		
2006	6166318	4741.6	953	2874.1	914.4		
2007	6338755	4502.6	935.4	2780.5	786.7		
2008	6500180	4087.3	894.2	2605.3	587.8		

Reported crime in Arkansas

Year	Population		Property crime rate		Burglary rate	Larceny-theft rate	Motor vehicle theft rate
2004	2750000	4033.1	1096.4	2699.7	237		
2005	2775708	4068	1085.1	2720	262		
2006	2810872	4021.6	1154.4	2596.7	270.4		
2007	2834797	3945.5	1124.4	2574.6	246.5		
2008	2855390	3843.7	1182.7	2433.4	227.6		

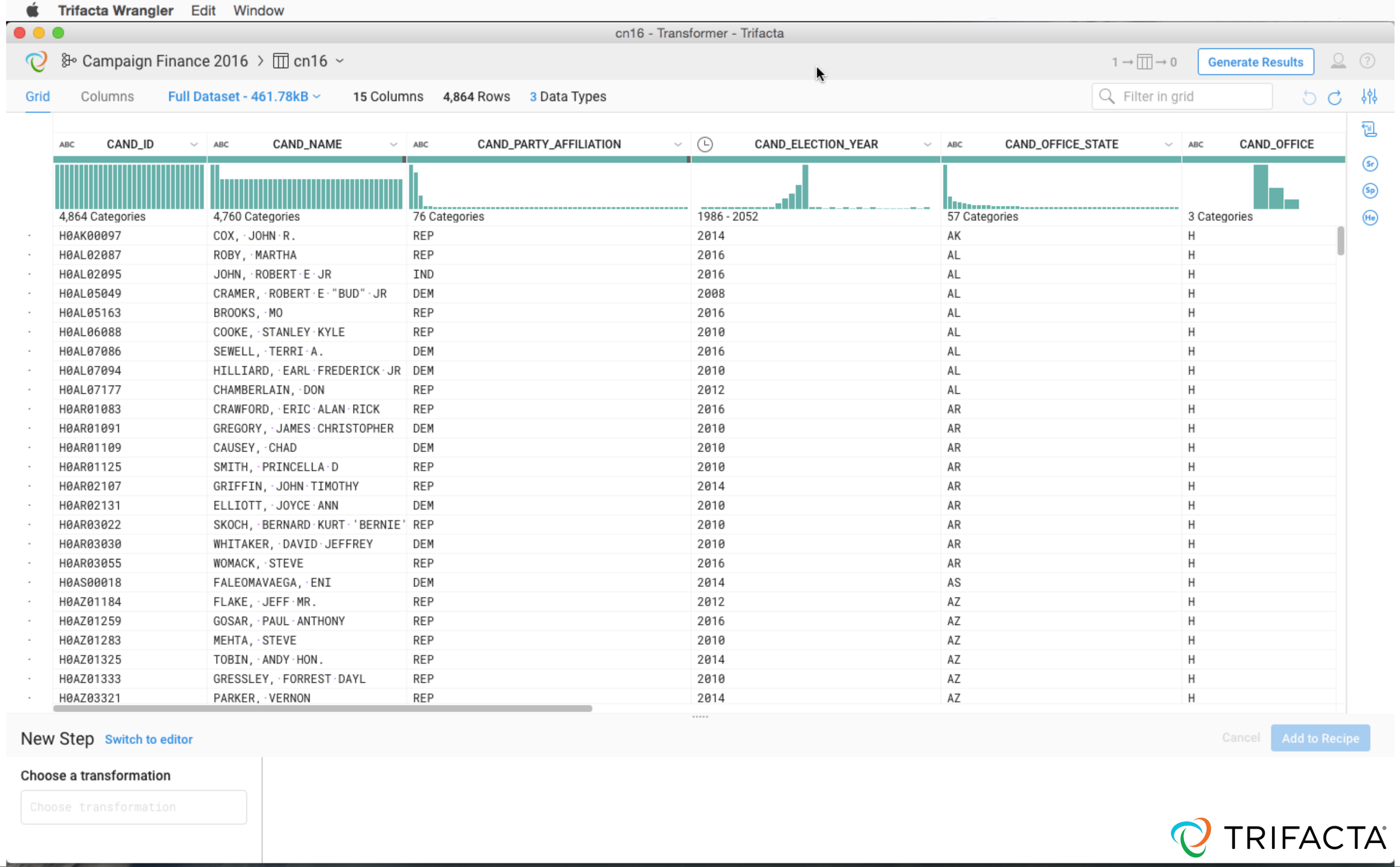
DataWrangler

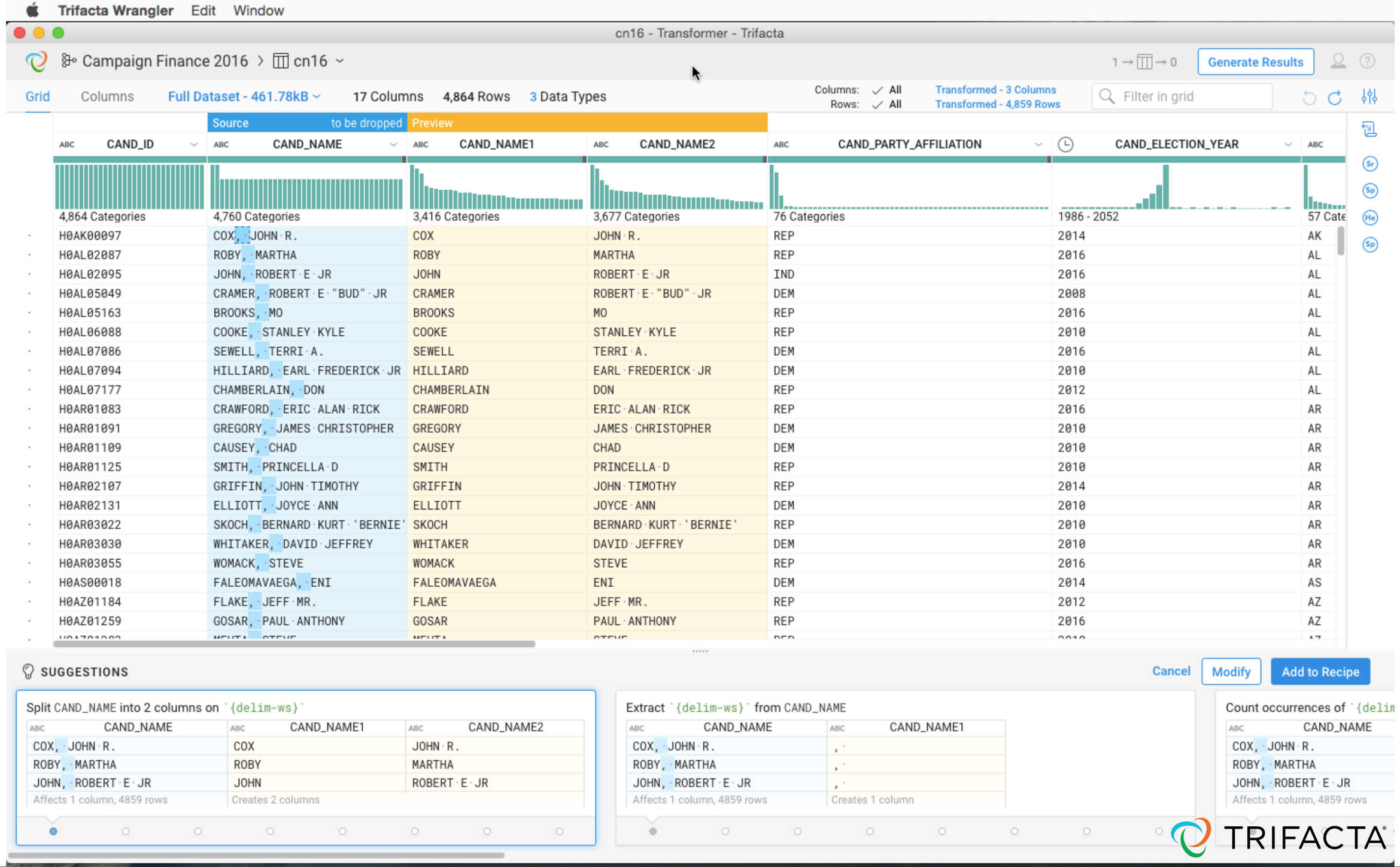


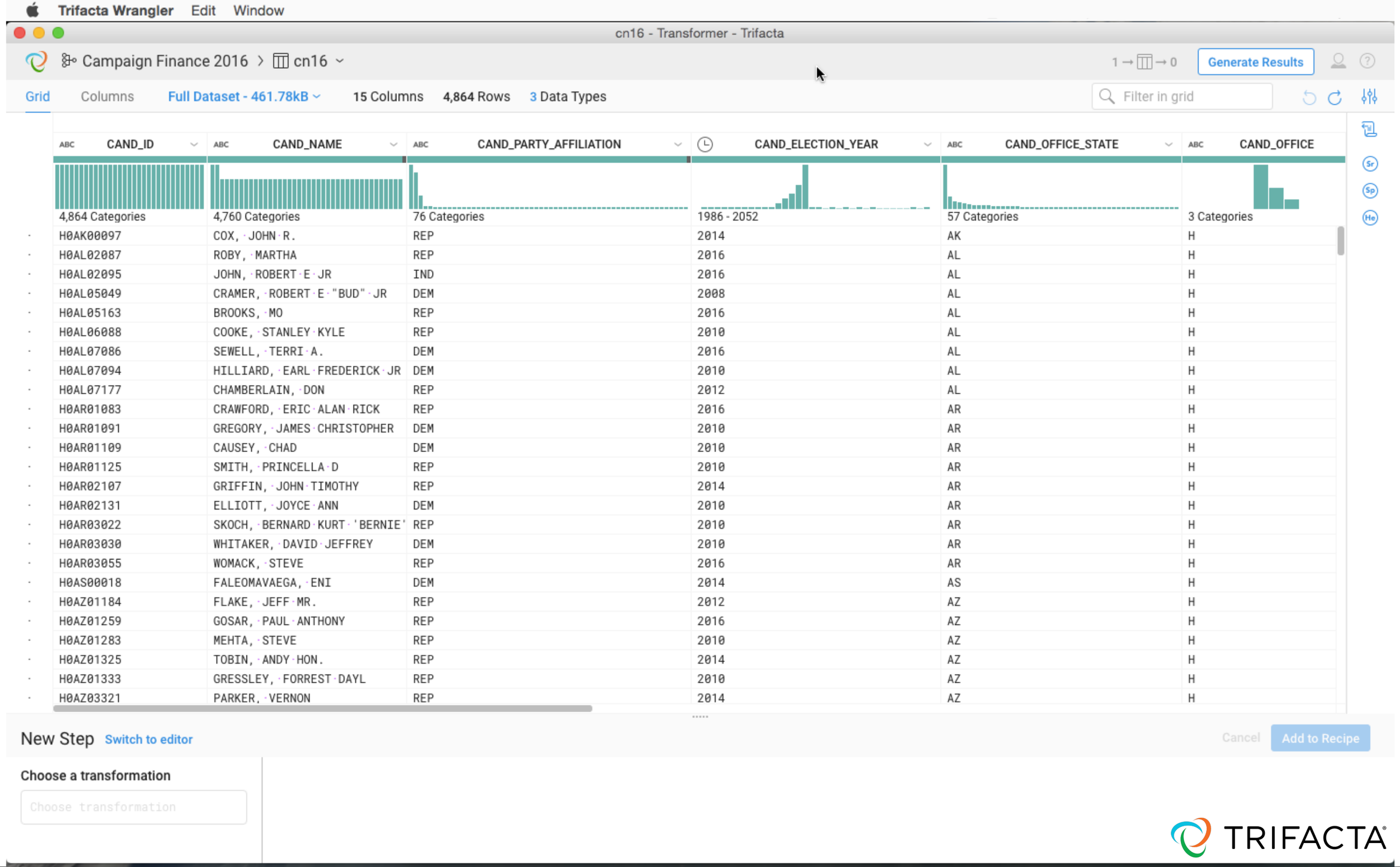
Wrangler: Interactive Visual Specification of Data Transformation Scripts

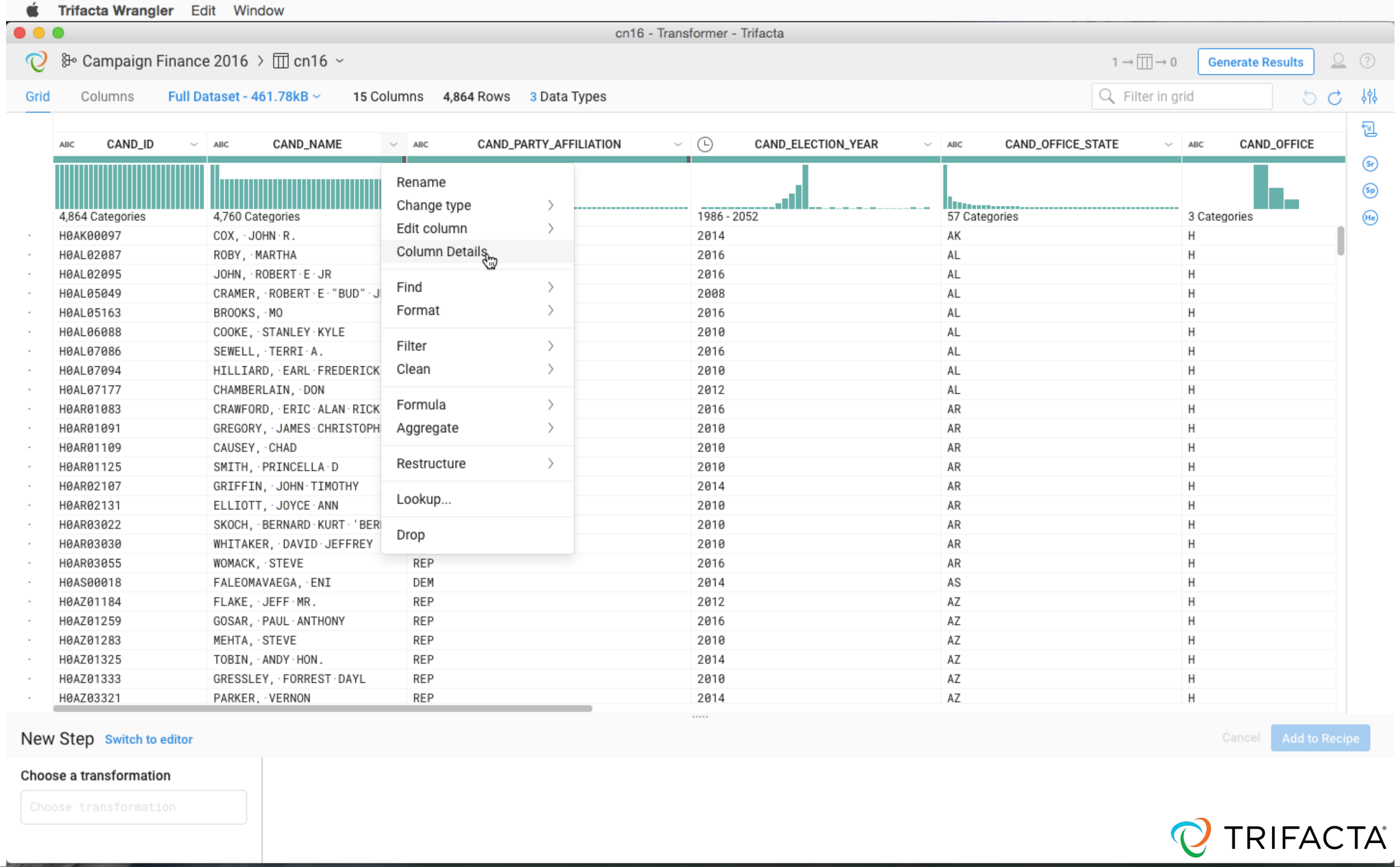
Sean Kandel, Joseph M. Hellerstein, et al. *CHI'11*

TRIFACTA®


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → 0
Generate Results
Grid Columns Full Dataset - 461.78kB 15 Columns 4,864 Rows 3 Data Types
Filter in grid
ABC CAND_ID
ABC CAND_NAME
ABC CAND_PARTY_AFFILIATION
CAND_ELECTION_YEAR
ABC CAND_OFFICE_STATE
ABC CAND_OFFICE
4,864 Categories
4,760 Categories
76 Categories
1986 - 2052
57 Categories
3 Categories
H0AK00097 COX, JOHN R. REP 2014 AK H
H0AL02087 ROBY, MARTHA REP 2016 AL H
H0AL02095 JOHN, ROBERT E JR IND 2016 AL H
H0AL05049 CRAMER, ROBERT E "BUD" JR DEM 2008 AL H
H0AL05163 BROOKS, MO REP 2016 AL H
H0AL06088 COOKE, STANLEY KYLE REP 2010 AL H
H0AL07086 SEWELL, TERRI A. DEM 2016 AL H
H0AL07094 HILLIARD, EARL FREDERICK JR DEM 2010 AL H
H0AL07177 CHAMBERLAIN, DON REP 2012 AL H
H0AR01083 CRAWFORD, ERIC ALAN RICK REP 2016 AR H
H0AR01091 GREGORY, JAMES CHRISTOPHER DEM 2010 AR H
H0AR01109 CAUSEY, CHAD DEM 2010 AR H
H0AR01125 SMITH, PRINCELLA D REP 2010 AR H
H0AR02107 GRIFFIN, JOHN TIMOTHY REP 2014 AR H
H0AR02131 ELLIOTT, JOYCE ANN DEM 2010 AR H
H0AR03022 SKOCH, BERNARD KURT 'BERNIE' REP 2010 AR H
H0AR03030 WHITAKER, DAVID JEFFREY DEM 2010 AR H
H0AR03055 WOMACK, STEVE REP 2016 AR H
H0AS00018 FALEOMAVAEGA, ENI DEM 2014 AS H
H0AZ01184 FLAKE, JEFF MR. REP 2012 AZ H
H0AZ01259 GOSAR, PAUL ANTHONY REP 2016 AZ H
H0AZ01283 MEHTA, STEVE REP 2010 AZ H
H0AZ01325 TOBIN, ANDY HON. REP 2014 AZ H
H0AZ01333 GRESSLEY, FORREST DAYL REP 2010 AZ H
H0AZ03321 PARKER, VERNON REP 2014 AZ H
New Step Switch to editor
Cancel
Add to Recipe
Choose a transformation
Choose transformation
TRIFACTA


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
Generate Results
Grid Columns Full Dataset - 461.78kB 17 Columns 4,864 Rows 3 Data Types
Columns: All Rows: All
Transformed - 3 Columns
Transformed - 4,859 Rows
Filter in grid
Source to be dropped
Preview
CAND_ID CAND_NAME CAND_NAME1 CAND_NAME2 CAND_PARTY_AFFILIATION CAND_ELECTION_YEAR
4,864 Categories 4,760 Categories 3,416 Categories 3,677 Categories 76 Categories 1986 - 2052 57 Cate
H0AK00097 COX, JOHN R. COX JOHN R. REP 2014 AK
H0AL02087 ROBY, MARTHA ROBY MARTHA REP 2016 AL
H0AL02095 JOHN, ROBERT E JR JOHN ROBERT E JR IND 2016 AL
H0AL05049 CRAMER, ROBERT E "BUD" JR CRAMER ROBERT E "BUD" JR DEM 2008 AL
H0AL05163 BROOKS, MO BROOKS MO REP 2016 AL
H0AL06088 COOKE, STANLEY KYLE COOKE STANLEY KYLE REP 2010 AL
H0AL07086 SEWELL, TERRI A. SEWELL TERRI A. DEM 2016 AL
H0AL07094 HILLIARD, EARL FREDERICK JR HILLIARD EARL FREDERICK JR DEM 2010 AL
H0AL07177 CHAMBERLAIN, DON CHAMBERLAIN DON REP 2012 AL
H0AR01083 CRAWFORD, ERIC ALAN RICK CRAWFORD ERIC ALAN RICK REP 2016 AR
H0AR01091 GREGORY, JAMES CHRISTOPHER GREGORY JAMES CHRISTOPHER DEM 2010 AR
H0AR01109 CAUSEY, CHAD CAUSEY CHAD DEM 2010 AR
H0AR01125 SMITH, PRINCELLA D SMITH PRINCELLA D REP 2010 AR
H0AR02107 GRIFFIN, JOHN TIMOTHY GRIFFIN JOHN TIMOTHY REP 2014 AR
H0AR02131 ELLIOTT, JOYCE ANN ELLIOTT JOYCE ANN DEM 2010 AR
H0AR03022 SKOCH, BERNARD KURT 'BERNIE' SKOCH BERNARD KURT 'BERNIE' REP 2010 AR
H0AR03030 WHITAKER, DAVID JEFFREY WHITAKER DAVID JEFFREY DEM 2010 AR
H0AR03055 WOMACK, STEVE WOMACK STEVE REP 2016 AR
H0AS00018 FALEOMAVAEGA, ENI FALEOMAVAEGA ENI DEM 2014 AS
H0AZ01184 FLAKE, JEFF MR. FLAKE JEFF MR. REP 2012 AZ
H0AZ01259 GOSAR, PAUL ANTHONY GOSAR PAUL ANTHONY REP 2016 AZ
SUGGESTIONS
Cancel
Modify
Add to Recipe
Split CAND_NAME into 2 columns on `{delim-ws}`
CAND_NAME CAND_NAME1 CAND_NAME2
COX, JOHN R. COX JOHN R.
ROBY, MARTHA ROBY MARTHA
JOHN, ROBERT E JR JOHN ROBERT E JR
Affects 1 column, 4859 rows
Creates 2 columns
Extract `{delim-ws}` from CAND_NAME
CAND_NAME CAND_NAME1
COX, JOHN R. ,
ROBY, MARTHA ,
JOHN, ROBERT E JR ,
Affects 1 column, 4859 rows
Creates 1 column
Count occurrences of `{delim
CAND_NAME
COX, JOHN R.
ROBY, MARTHA
JOHN, ROBERT E JR
Affects 1 column, 4859 rows
TRIFACTA


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → 0
Generate Results
Grid Columns Full Dataset - 461.78kB 15 Columns 4,864 Rows 3 Data Types
Filter in grid
ABC CAND_ID
ABC CAND_NAME
ABC CAND_PARTY_AFFILIATION
CAND_ELECTION_YEAR
ABC CAND_OFFICE_STATE
ABC CAND_OFFICE
4,864 Categories
4,760 Categories
76 Categories
1986 - 2052
57 Categories
3 Categories
H0AK00097 COX, JOHN R. REP 2014 AK H
H0AL02087 ROBY, MARTHA REP 2016 AL H
H0AL02095 JOHN, ROBERT E JR IND 2016 AL H
H0AL05049 CRAMER, ROBERT E "BUD" JR DEM 2008 AL H
H0AL05163 BROOKS, MO REP 2016 AL H
H0AL06088 COOKE, STANLEY KYLE REP 2010 AL H
H0AL07086 SEWELL, TERRI A. DEM 2016 AL H
H0AL07094 HILLIARD, EARL FREDERICK JR DEM 2010 AL H
H0AL07177 CHAMBERLAIN, DON REP 2012 AL H
H0AR01083 CRAWFORD, ERIC ALAN RICK REP 2016 AR H
H0AR01091 GREGORY, JAMES CHRISTOPHER DEM 2010 AR H
H0AR01109 CAUSEY, CHAD DEM 2010 AR H
H0AR01125 SMITH, PRINCELLA D REP 2010 AR H
H0AR02107 GRIFFIN, JOHN TIMOTHY REP 2014 AR H
H0AR02131 ELLIOTT, JOYCE ANN DEM 2010 AR H
H0AR03022 SKOCH, BERNARD KURT 'BERNIE' REP 2010 AR H
H0AR03030 WHITAKER, DAVID JEFFREY DEM 2010 AR H
H0AR03055 WOMACK, STEVE REP 2016 AR H
H0AS00018 FALEOMAVAEGA, ENI DEM 2014 AS H
H0AZ01184 FLAKE, JEFF MR. REP 2012 AZ H
H0AZ01259 GOSAR, PAUL ANTHONY REP 2016 AZ H
H0AZ01283 MEHTA, STEVE REP 2010 AZ H
H0AZ01325 TOBIN, ANDY HON. REP 2014 AZ H
H0AZ01333 GRESSLEY, FORREST DAYL REP 2010 AZ H
H0AZ03321 PARKER, VERNON REP 2014 AZ H
New Step Switch to editor
Cancel
Add to Recipe
Choose a transformation
Choose transformation
TRIFACTA


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → 0
Generate Results
Grid
Columns
Full Dataset - 461.78kB
15 Columns
4,864 Rows
3 Data Types
Filter in grid
ABC CAND_ID
ABC CAND_NAME
ABC CAND_PARTY_AFFILIATION
CAND_ELECTION_YEAR
ABC CAND_OFFICE_STATE
ABC CAND_OFFICE
4,864 Categories
4,760 Categories
1986 - 2052
57 Categories
3 Categories
H0AK00097 COX, JOHN R. 2014 AK H
H0AL02087 ROBY, MARTHA 2016 AL H
H0AL02095 JOHN, ROBERT E JR 2016 AL H
H0AL05049 CRAMER, ROBERT E "BUD" J 2008 AL H
H0AL05163 BROOKS, MO 2016 AL H
H0AL06088 COOKE, STANLEY KYLE 2010 AL H
H0AL07086 SEWELL, TERRI A. 2016 AL H
H0AL07094 HILLIARD, EARL FREDERICK 2010 AL H
H0AL07177 CHAMBERLAIN, DON 2012 AL H
H0AR01083 CRAWFORD, ERIC ALAN RICK 2016 AR H
H0AR01091 GREGORY, JAMES CHRISTOPH 2010 AR H
H0AR01109 CAUSEY, CHAD 2010 AR H
H0AR01125 SMITH, PRINCELLA D 2010 AR H
H0AR02107 GRIFFIN, JOHN TIMOTHY 2014 AR H
H0AR02131 ELLIOTT, JOYCE ANN 2010 AR H
H0AR03022 SKOCH, BERNARD KURT 'BER 2010 AR H
H0AR03030 WHITAKER, DAVID JEFFREY 2010 AR H
H0AR03055 WOMACK, STEVE REP 2016 AR H
H0AS00018 FALEOMAVAEGA, ENI DEM 2014 AS H
H0AZ01184 FLAKE, JEFF MR. REP 2012 AZ H
H0AZ01259 GOSAR, PAUL ANTHONY REP 2016 AZ H
H0AZ01283 MEHTA, STEVE REP 2010 AZ H
H0AZ01325 TOBIN, ANDY HON. REP 2014 AZ H
H0AZ01333 GRESSLEY, FORREST DAYL REP 2010 AZ H
H0AZ03321 PARKER, VERNON REP 2014 AZ H
Rename
Change type
Edit column
Column Details
Find
Format
Filter
Clean
Formula
Aggregate
Restructure
Lookup...
Drop
New Step
Switch to editor
Cancel
Add to Recipe
Choose a transformation
Choose transformation
TRIFACTA

Trifacta Wrangler Edit Window

cn16 - Transformer - Trifacta

Campaign Finance 2016 › cn16

1 → → 0 Generate Results

Grid Columns Full Dataset - 461.78kB 15 Columns 4,864 Rows 3 Data Types

CAND_NAME

SUMMARY

Valid	4,863	100.0%
Unique	4,760	97.9%
Outliers	18	0.4%
Mismatched	0	0.0%
Missing	1	0.0%

STRING LENGTH STATISTICS

Minimum	4.00
Lower Quartile	14.00
Median	18.00
Upper Quartile	21.00
Maximum	70.00
Average	18.14
Standard Deviation	4.99

TOP VALUES

KALEMKARIAN, TIMOTHY CHARLES	3
MARTIN, ANDY	3
AGBEDE, AKINYEMI	2
AKIN, W TODD	2
ARMSTRONG, BRANDON CHRISTINA	2
BACHMANN, MICHELE	2
BALDWIN, TAMMY	2
BARR, BOB	2
BATES, DON JR	2
BELLIS, JOSEPH K III	2
BICKELMEYER, MICHAEL	2
BLASS, PIOTR DR	2
BLUNT, ROY	2
BOSS, JEFF	2

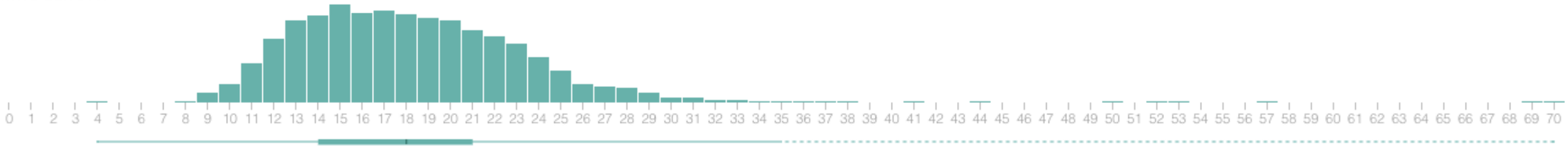
MISMATCHED VALUES

None

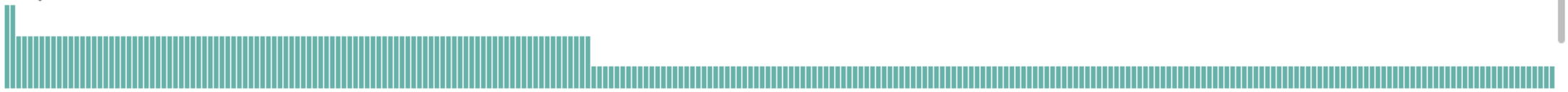
STRING LENGTH OUTLIERS

AAAAAAAAAAAAAAAAAAAAAAAAAAAAA…	1
AKA THE PROPHET AKA EARL, TRIP…	1
CLARKSON, JEREMY CHARLES ROBER…	1
CONNOLLY, MATTHEW DONALD (MATT…	1
DE BUONAPARTE, HRM CAESAR ST A…	1
EASTON, EARNEST LEE PROFESSOR …	1

STRING LENGTH



FREQUENT VALUES



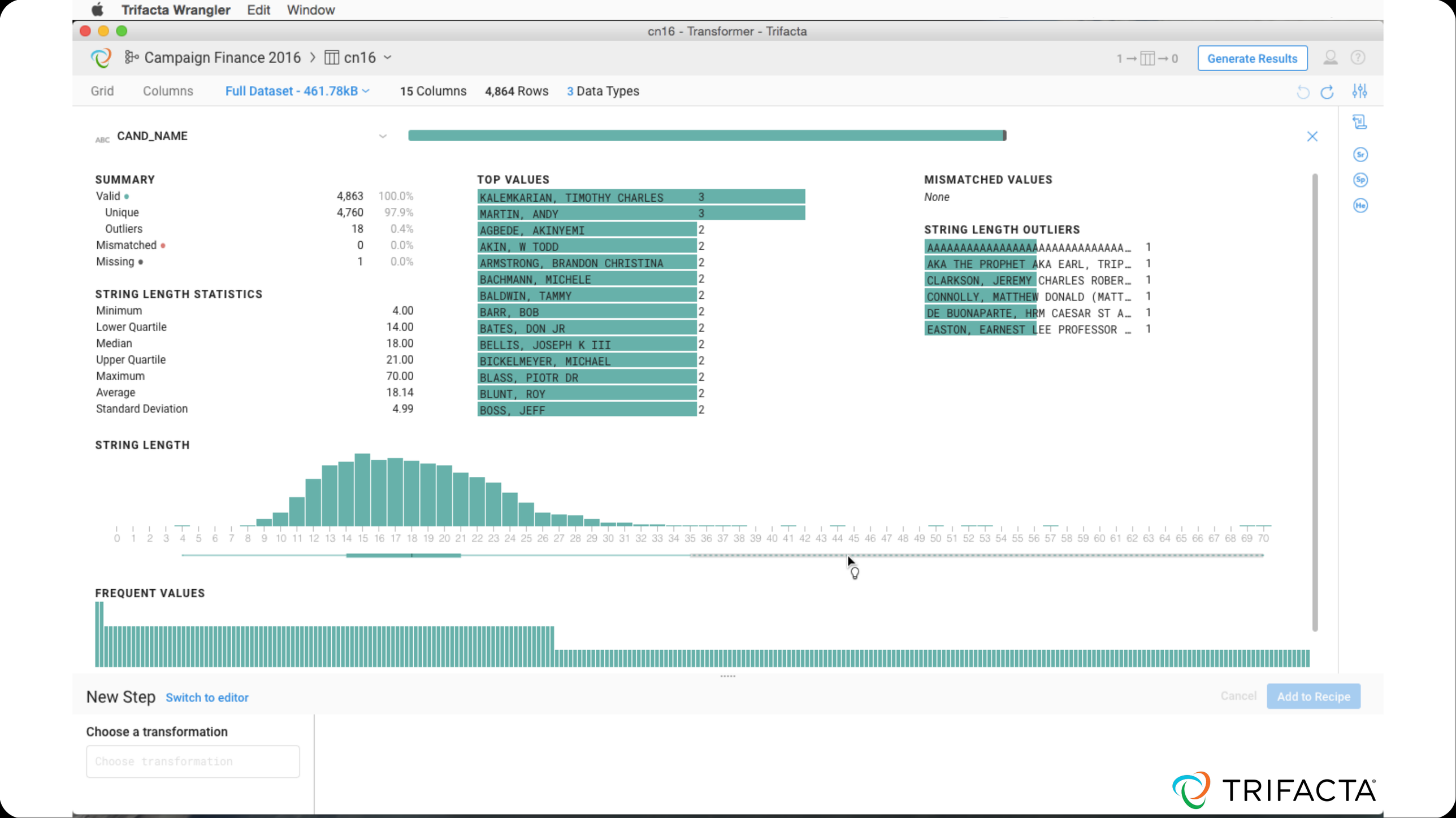
New Step Switch to editor

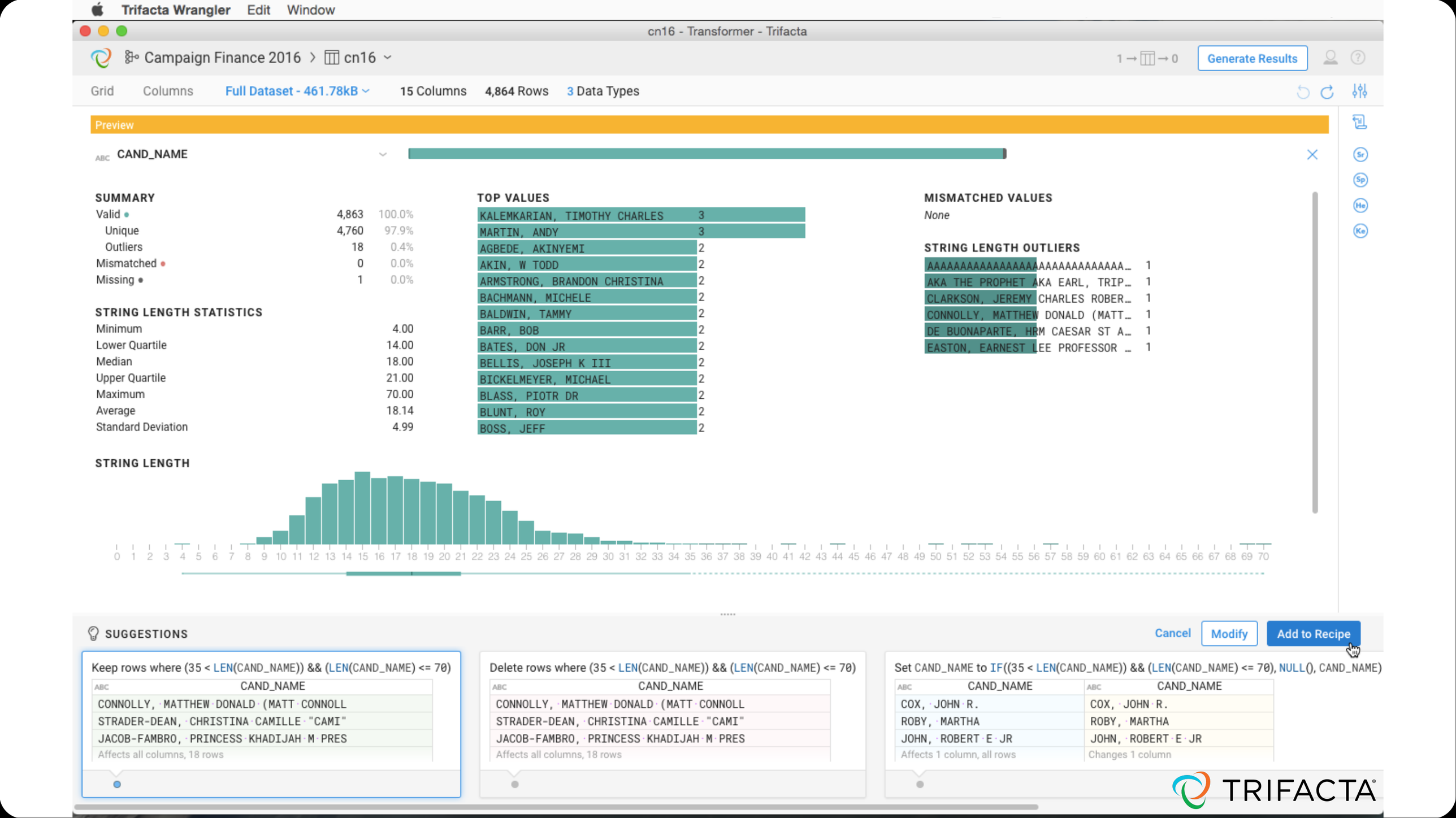
Cancel Add to Recipe

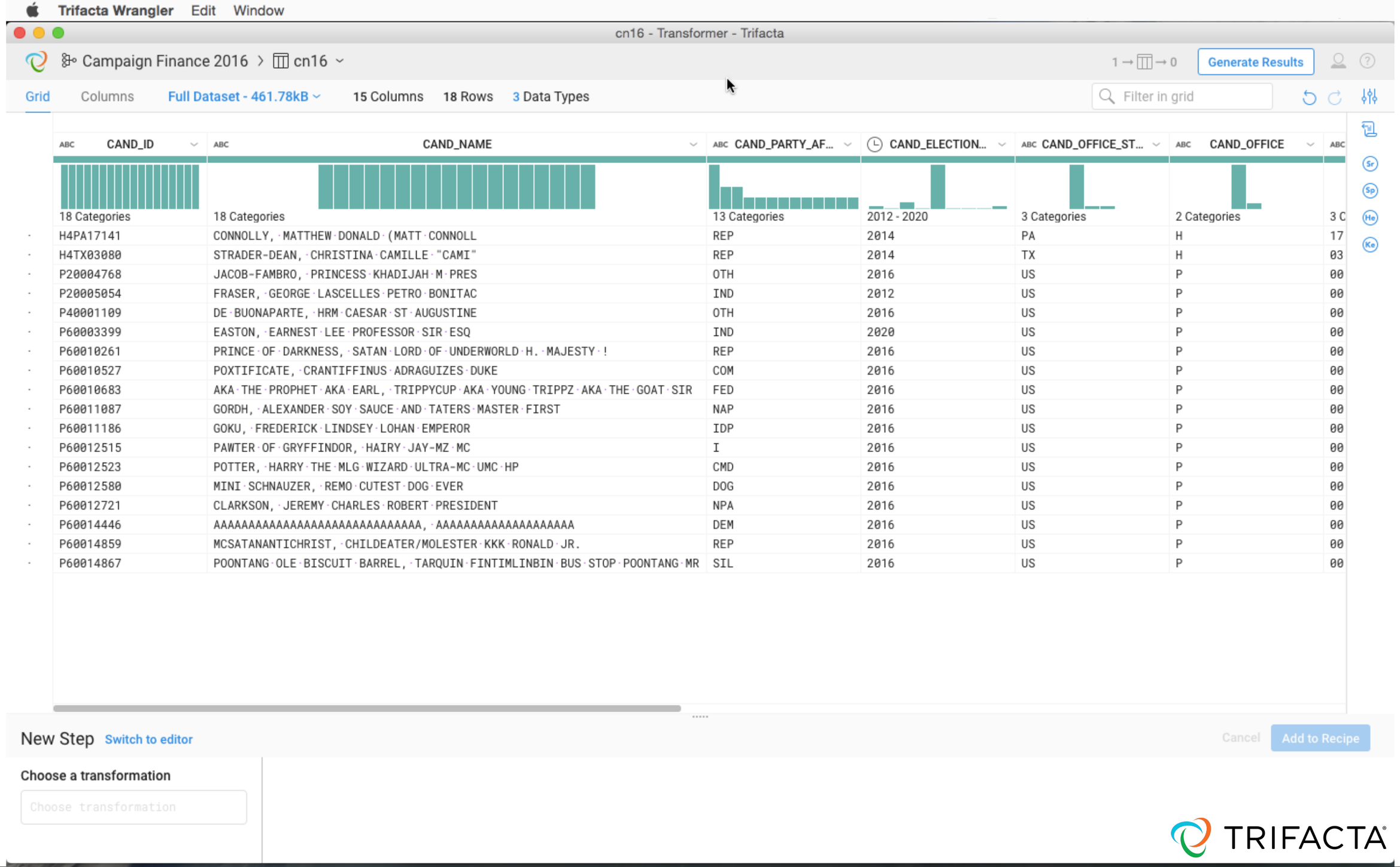
Choose a transformation

Choose transformation

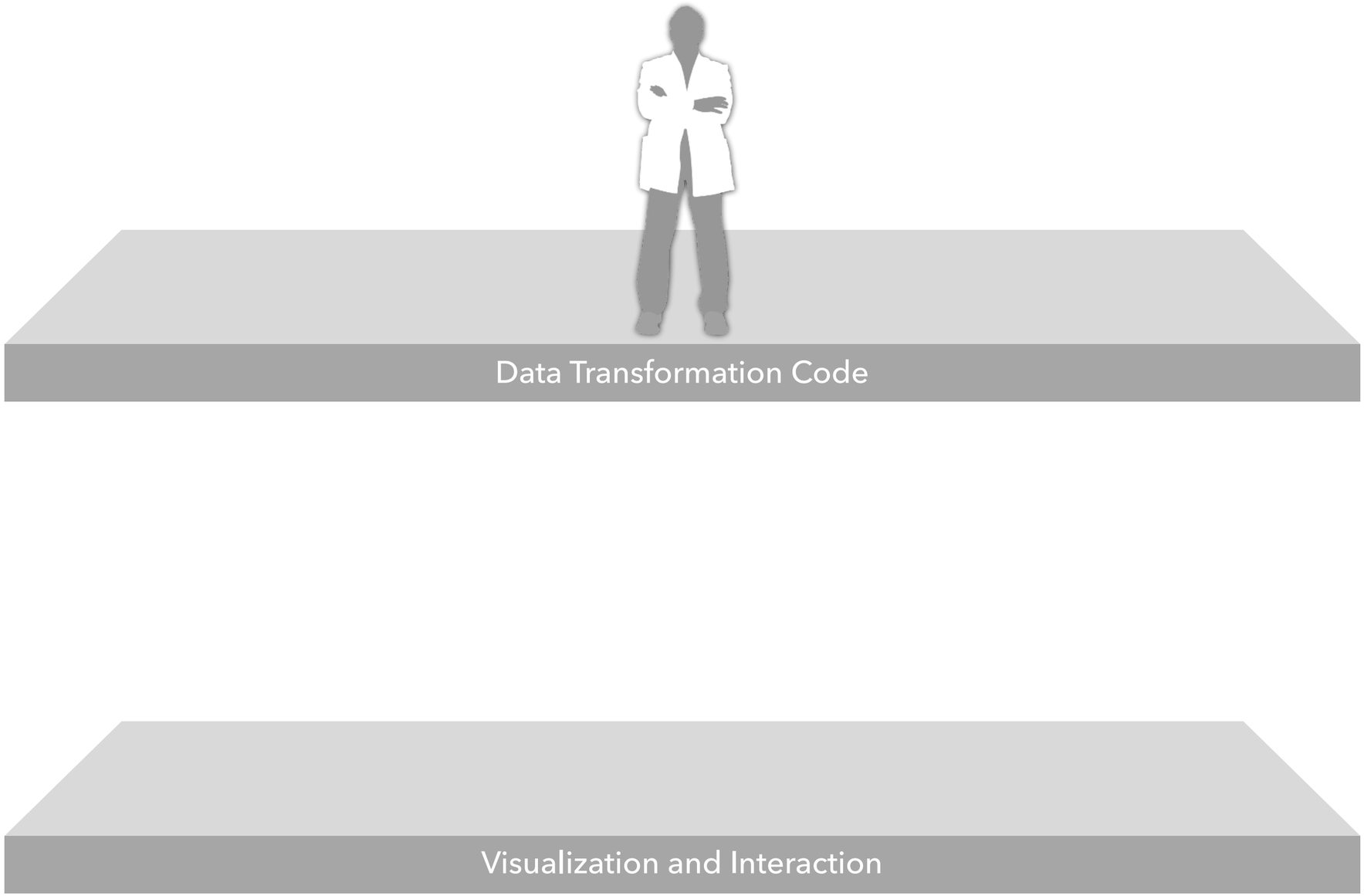
TRIFACTA


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → → 0
Generate Results
Grid Columns Full Dataset - 461.78kB 15 Columns 4,864 Rows 3 Data Types
ABC CAND_NAME
SUMMARY
Valid 4,863 100.0%
Unique 4,760 97.9%
Outliers 18 0.4%
Mismatched 0 0.0%
Missing 1 0.0%
STRING LENGTH STATISTICS
Minimum 4.00
Lower Quartile 14.00
Median 18.00
Upper Quartile 21.00
Maximum 70.00
Average 18.14
Standard Deviation 4.99
TOP VALUES
KALEMKARIAN, TIMOTHY CHARLES 3
MARTIN, ANDY 3
AGBEDE, AKINYEMI 2
AKIN, W TODD 2
ARMSTRONG, BRANDON CHRISTINA 2
BACHMANN, MICHELE 2
BALDWIN, TAMMY 2
BARR, BOB 2
BATES, DON JR 2
BELLIS, JOSEPH K III 2
BICKELMEYER, MICHAEL 2
BLASS, PIOTR DR 2
BLUNT, ROY 2
BOSS, JEFF 2
MISMATCHED VALUES
None
STRING LENGTH OUTLIERS
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA… 1
AKA THE PROPHET AKA EARL, TRIP… 1
CLARKSON, JEREMY CHARLES ROBER… 1
CONNOLLY, MATTHEW DONALD (MATT… 1
DE BUONAPARTE, HRM CAESAR ST A… 1
EASTON, EARNEST LEE PROFESSOR … 1
STRING LENGTH
0 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 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70
FREQUENT VALUES
New Step Switch to editor
Cancel Add to Recipe
Choose a transformation
Choose transformation
TRIFACTA


Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → → 0
Generate Results
Grid Columns Full Dataset - 461.78kB 15 Columns 4,864 Rows 3 Data Types
Preview
CAND_NAME
SUMMARY
Valid 4,863 100.0%
Unique 4,760 97.9%
Outliers 18 0.4%
Mismatched 0 0.0%
Missing 1 0.0%
STRING LENGTH STATISTICS
Minimum 4.00
Lower Quartile 14.00
Median 18.00
Upper Quartile 21.00
Maximum 70.00
Average 18.14
Standard Deviation 4.99
TOP VALUES
KALEMKARIAN, TIMOTHY CHARLES 3
MARTIN, ANDY 3
AGBEDE, AKINYEMI 2
AKIN, W TODD 2
ARMSTRONG, BRANDON CHRISTINA 2
BACHMANN, MICHELE 2
BALDWIN, TAMMY 2
BARR, BOB 2
BATES, DON JR 2
BELLIS, JOSEPH K III 2
BICKELMEYER, MICHAEL 2
BLASS, PIOTR DR 2
BLUNT, ROY 2
BOSS, JEFF 2
MISMATCHED VALUES
None
STRING LENGTH OUTLIERS
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA… 1
AKA THE PROPHET AKA EARL, TRIP… 1
CLARKSON, JEREMY CHARLES ROBER… 1
CONNOLLY, MATTHEW DONALD (MATT… 1
DE BUONAPARTE, HRM CAESAR ST A… 1
EASTON, EARNEST LEE PROFESSOR … 1
STRING LENGTH
0 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 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70
SUGGESTIONS
Cancel
Modify
Add to Recipe
Keep rows where (35 < LEN(CAND_NAME)) && (LEN(CAND_NAME) <= 70)
CAND_NAME
CONNOLLY, MATTHEW DONALD (MATT CONNOLL
STRADER-DEAN, CHRISTINA CAMILLE "CAMI"
JACOB-FAMBRO, PRINCESS KHADIJAH M PRES
Affects all columns, 18 rows
Delete rows where (35 < LEN(CAND_NAME)) && (LEN(CAND_NAME) <= 70)
CAND_NAME
CONNOLLY, MATTHEW DONALD (MATT CONNOLL
STRADER-DEAN, CHRISTINA CAMILLE "CAMI"
JACOB-FAMBRO, PRINCESS KHADIJAH M PRES
Affects all columns, 18 rows
Set CAND_NAME to IF((35 < LEN(CAND_NAME)) && (LEN(CAND_NAME) <= 70), NULL(), CAND_NAME)
CAND_NAME
COX, JOHN R.
ROBY, MARTHA
JOHN, ROBERT E JR
Affects 1 column, all rows
CAND_NAME
COX, JOHN R.
ROBY, MARTHA
JOHN, ROBERT E JR
Changes 1 column
TRIFACTA

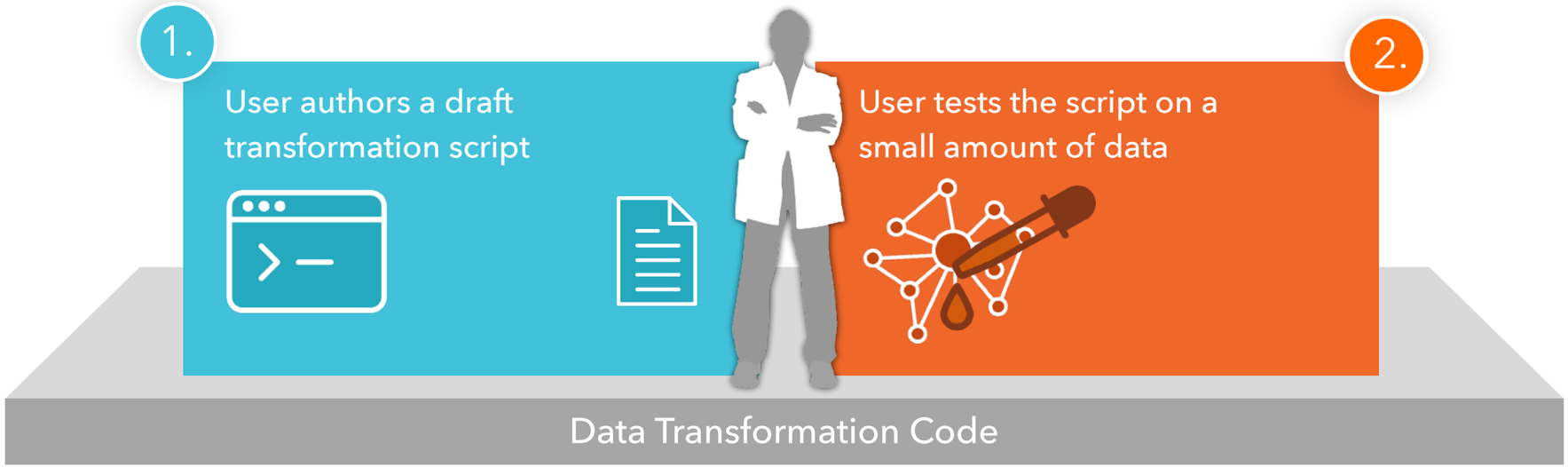

Trifacta Wrangler Edit Window
cn16 - Transformer - Trifacta
Campaign Finance 2016 › cn16
1 → 0
Generate Results
Grid
Columns
Full Dataset - 461.78kB
15 Columns
18 Rows
3 Data Types
Filter in grid
ABC CAND_ID
ABC CAND_NAME
ABC CAND_PARTY_AF...
CAND_ELECTION...
ABC CAND_OFFICE_ST...
ABC CAND_OFFICE
18 Categories
18 Categories
13 Categories
2012 - 2020
3 Categories
2 Categories
H4PA17141 CONNOLLY, MATTHEW DONALD (MATT CONNOLL REP 2014 PA H 17
H4TX03080 STRADER-DEAN, CHRISTINA CAMILLE "CAMI" REP 2014 TX H 03
P20004768 JACOB-FAMBRO, PRINCESS KHADIJAH M PRES OTH 2016 US P 00
P20005054 FRASER, GEORGE LASCELLES PETRO BONITAC IND 2012 US P 00
P40001109 DE BUONAPARTE, HRM CAESAR ST AUGUSTINE OTH 2016 US P 00
P60003399 EASTON, EARNEST LEE PROFESSOR SIR ESQ IND 2020 US P 00
P60010261 PRINCE OF DARKNESS, SATAN LORD OF UNDERWORLD H. MAJESTY ! REP 2016 US P 00
P60010527 POXTIFICATE, CRANTIFFINUS ADRAGUIZES DUKE COM 2016 US P 00
P60010683 AKA THE PROPHET AKA EARL, TRIPPYCUP AKA YOUNG TRIPPZ AKA THE GOAT SIR FED 2016 US P 00
P60011087 GORDH, ALEXANDER SOY SAUCE AND TATERS MASTER FIRST NAP 2016 US P 00
P60011186 GOKU, FREDERICK LINDSEY LOHAN EMPEROR IDP 2016 US P 00
P60012515 PAWTER OF GRYFFINDOR, HAIRY JAY-MZ MC I 2016 US P 00
P60012523 POTTER, HARRY THE MLG WIZARD ULTRA-MC UMC HP CMD 2016 US P 00
P60012580 MINI SCHNAUZER, REMO CUTEST DOG EVER DOG 2016 US P 00
P60012721 CLARKSON, JEREMY CHARLES ROBERT PRESIDENT NPA 2016 US P 00
P60014446 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA, AAAAAAAAAAAAAAAAAA DEM 2016 US P 00
P60014859 MCSATANANTICHRIST, CHILDEATER/MOLESTER KKK RONALD JR. REP 2016 US P 00
P60014867 POONTANG OLE BISCUIT BARREL, TARQUIN FINTIMLINBIN BUS STOP POONTANG MR SIL 2016 US P 00
New Step
Switch to editor
Cancel
Add to Recipe
Choose a transformation
Choose transformation
TRIFACTA

Traditional Specification

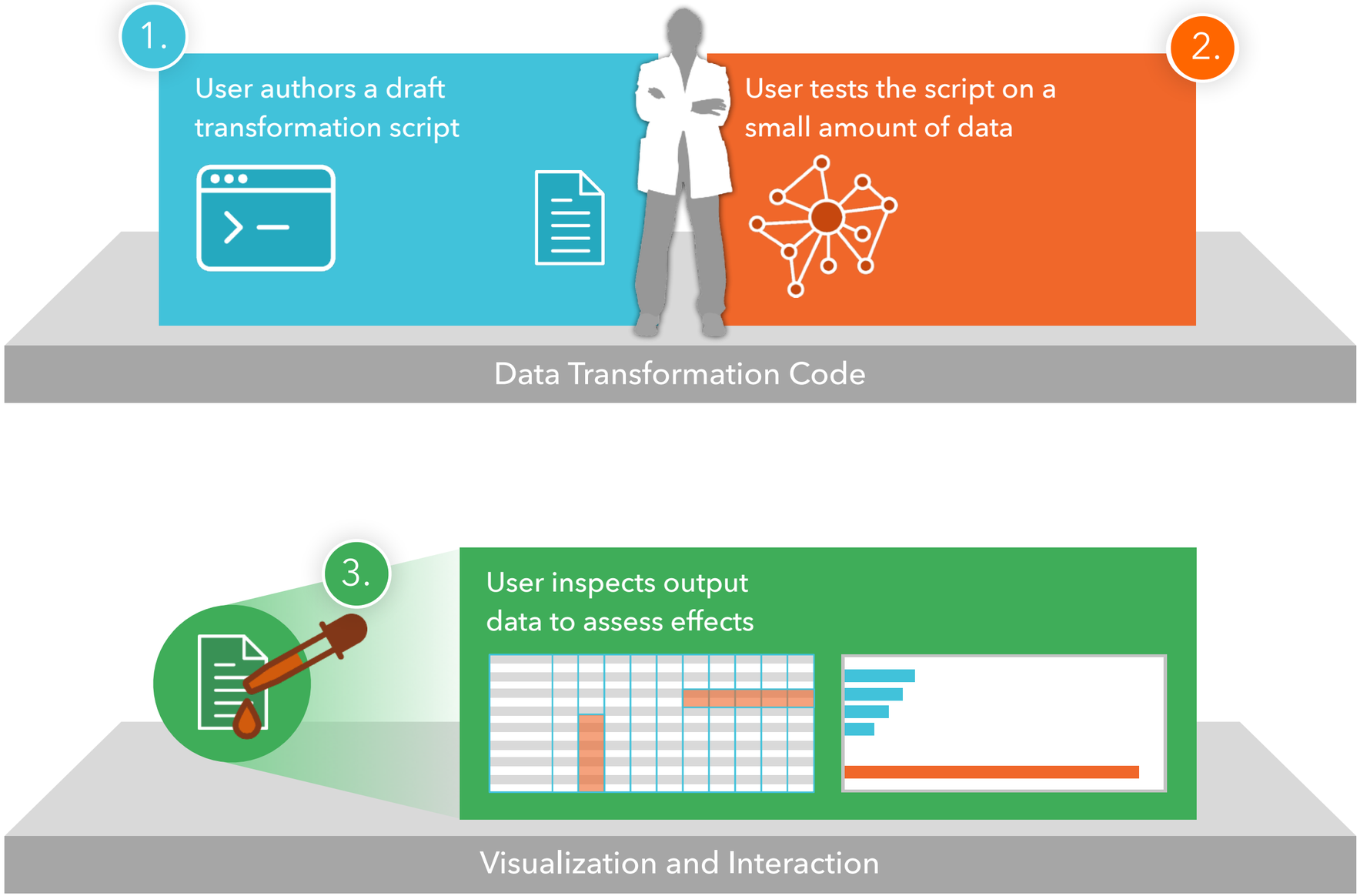


Traditional Specification

Traditional Specification

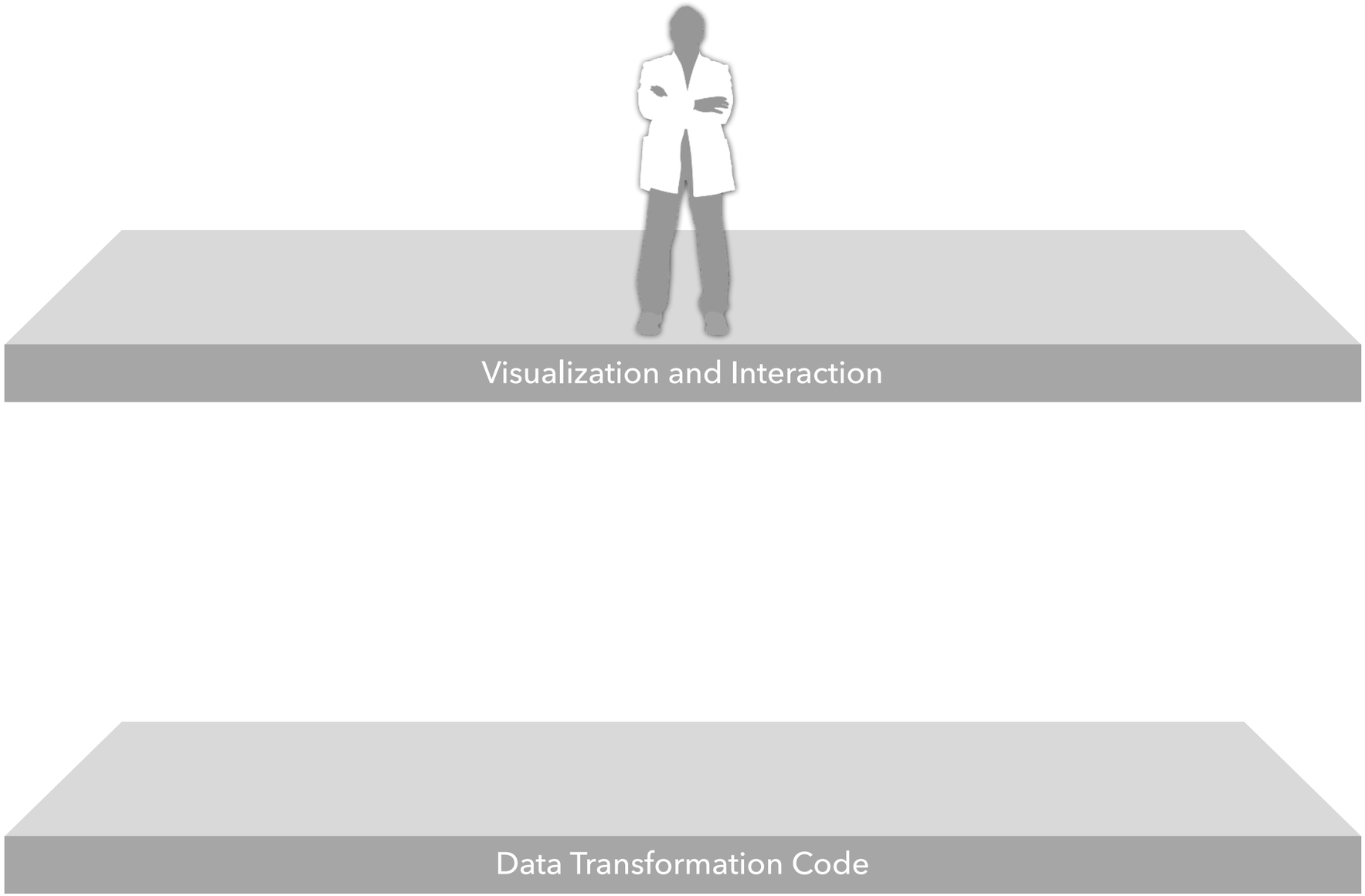


Traditional Specification

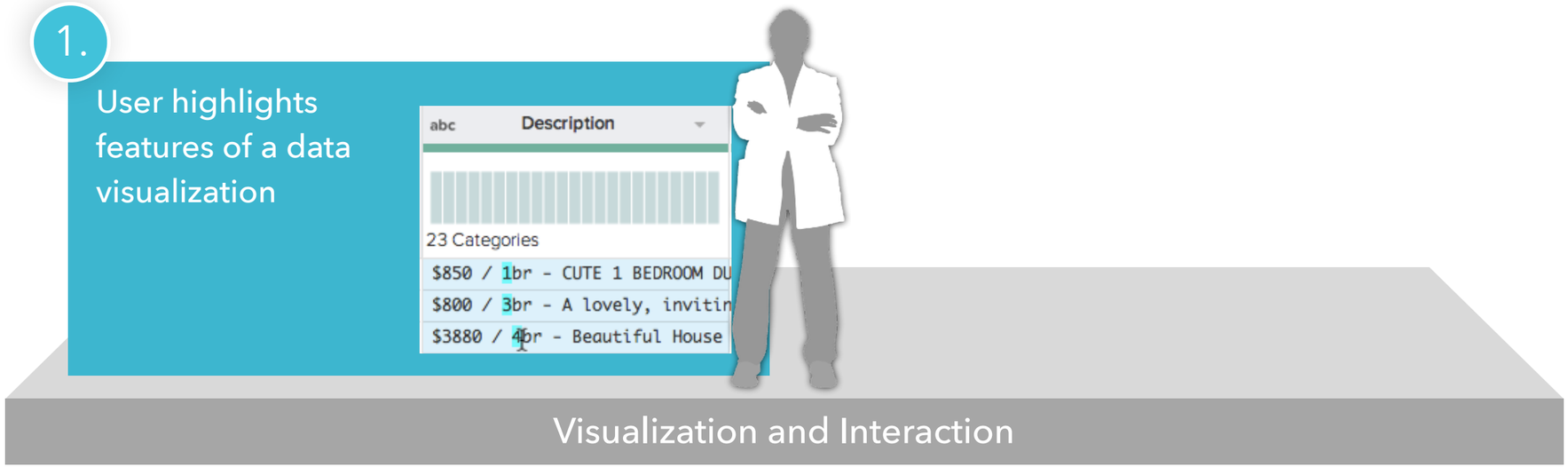


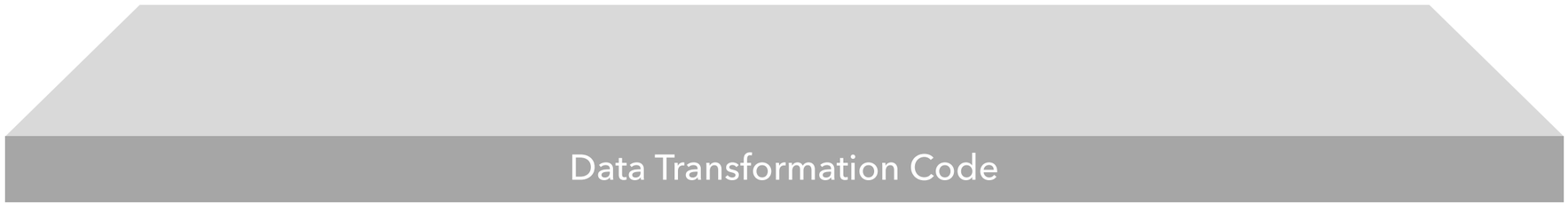
Predictive Interaction

Predictive Interaction

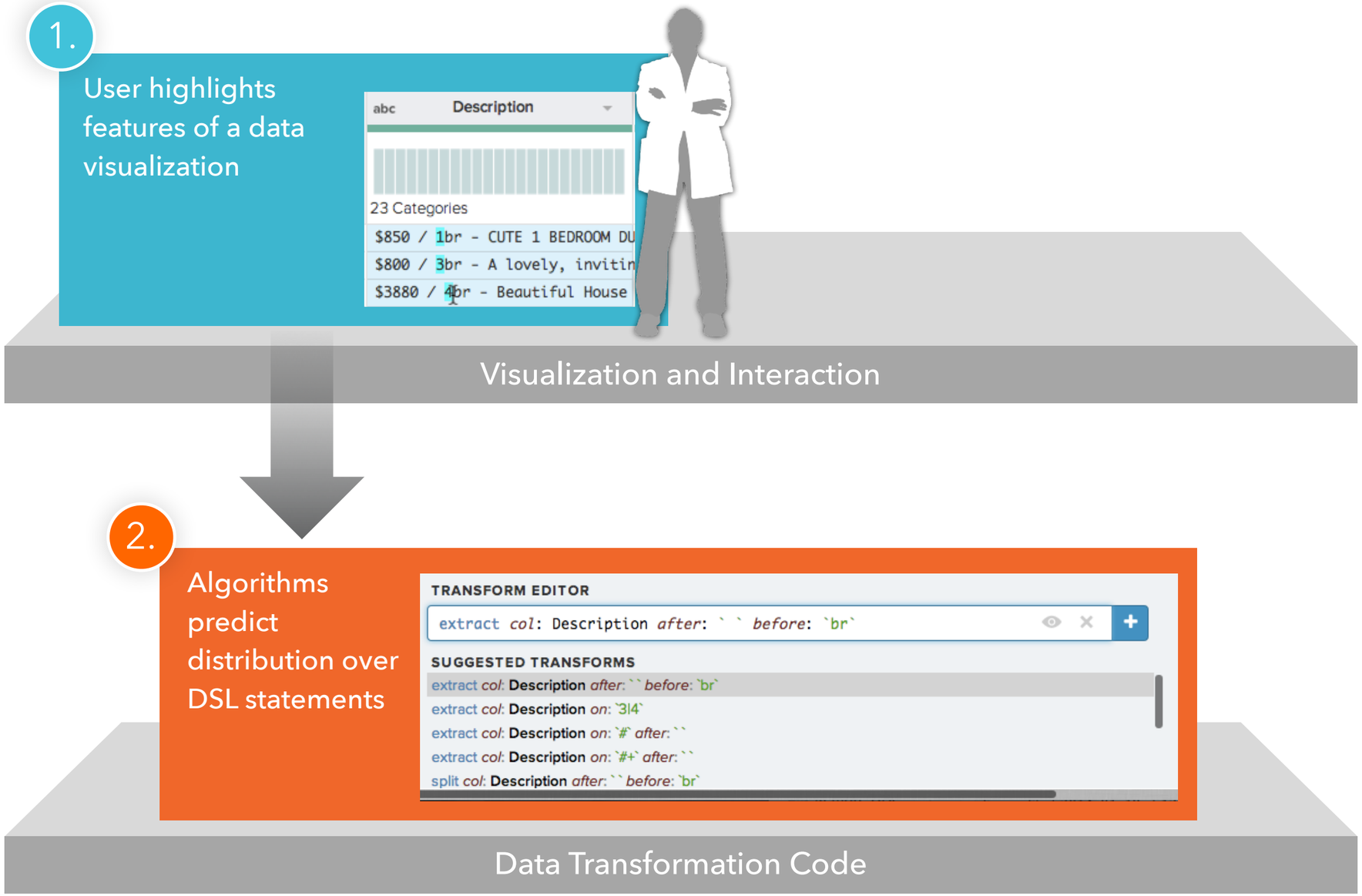


Predictive Interaction

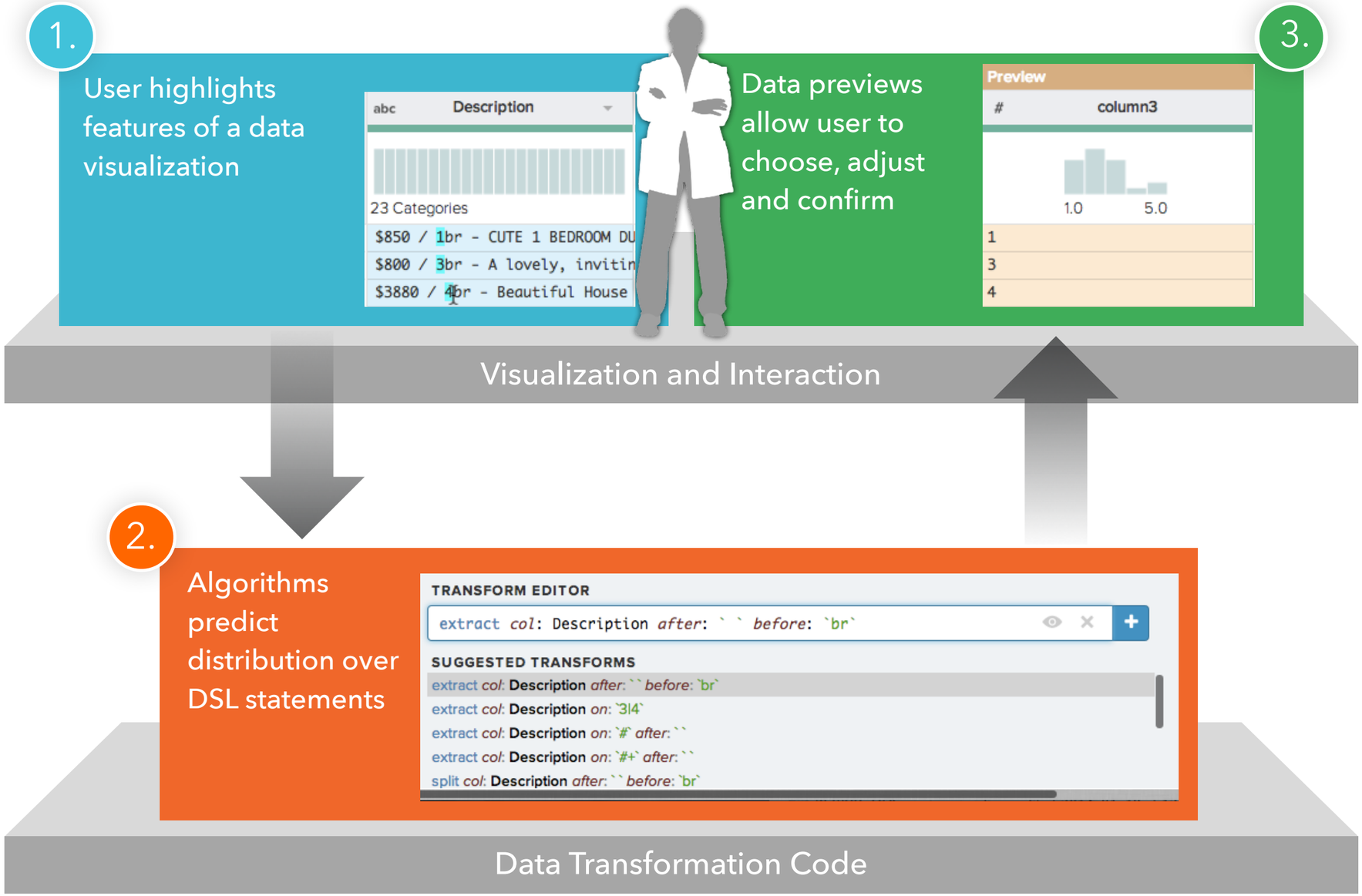




Predictive Interaction



Predictive Interaction



Wrangle Language Building Blocks

Transforms

Split

Extract

Filter

Derive

Header

Pivot

Aggregate

Join

Union

…

Wrangle Language Building Blocks

Transforms	Parameter Types
Split	Text Selection
Extract	Column Selection
Filter	Row Selection
Derive	Formula
Header	Enumeration
Pivot	Number
Aggregate	String
Join	Boolean
Union	
…	

Text Selection: A Language within a Language

Transforms	**Parameter Types**
Split	***Text Selection***
Extract	Column Selection
Filter	Row Selection
Derive	Formula
Header	Enumeration
Pivot	Number
Aggregate	String
Join	Boolean
Union	
…	

SUGGESTED TRANSFORMS

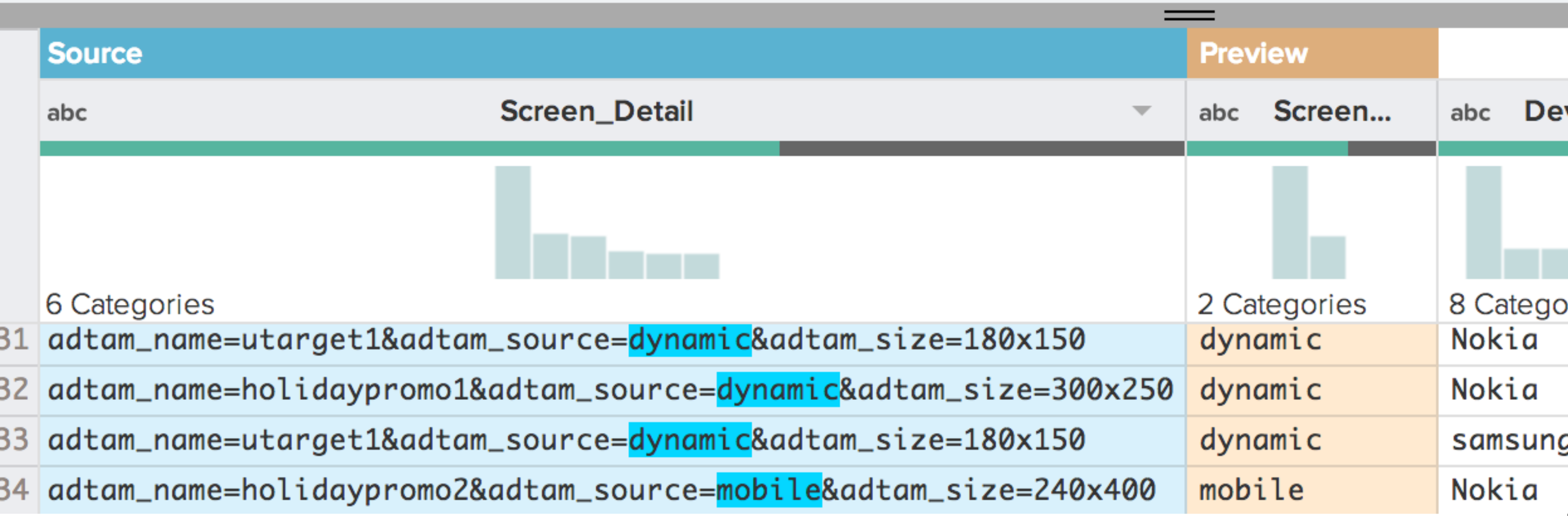
	Source	Preview	
	abc Screen_Detail	abc Screen...	abc De
	6 Categories	2 Categories	8 Catego
31	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	Nokia
32	adtam_name=holidaypromo1&adtam_source=dynamic&adtam_size=300x250	dynamic	Nokia
33	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	samsung
34	adtam_name=holidaypromo2&adtam_source=mobile&adtam_size=240x400	mobile	Nokia

SUGGESTED TRANSFORMS

extract *col:* **Screen_Detail** *on:* /(?<=adtam_source\=)[^\&]*(?=\&)/

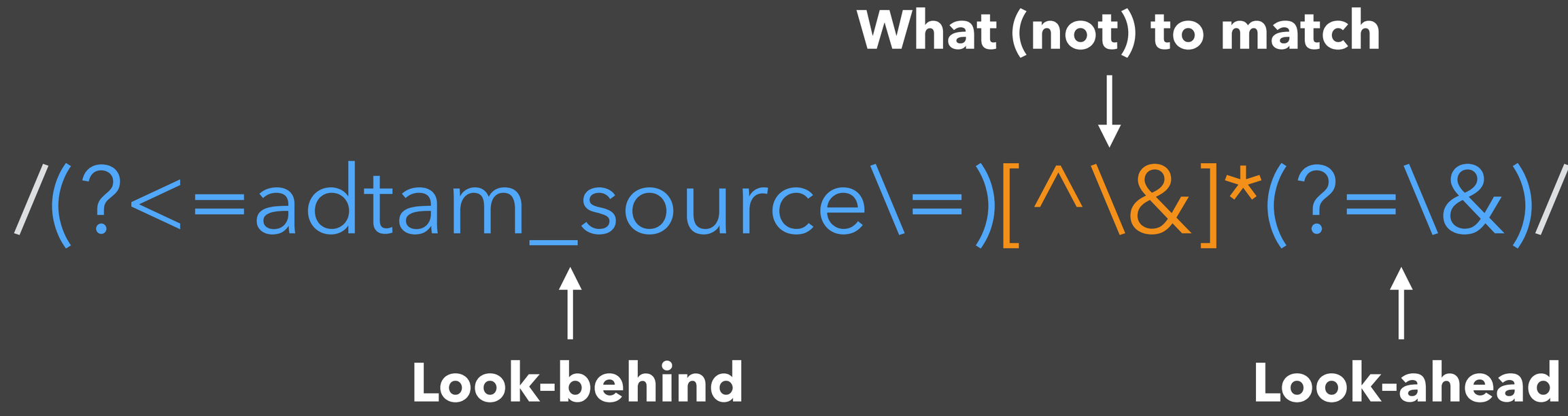
extract *col:* **Screen_Detail** *on:* /(?<=\=)[^\&]*(?=\&)/ *limit:* 2

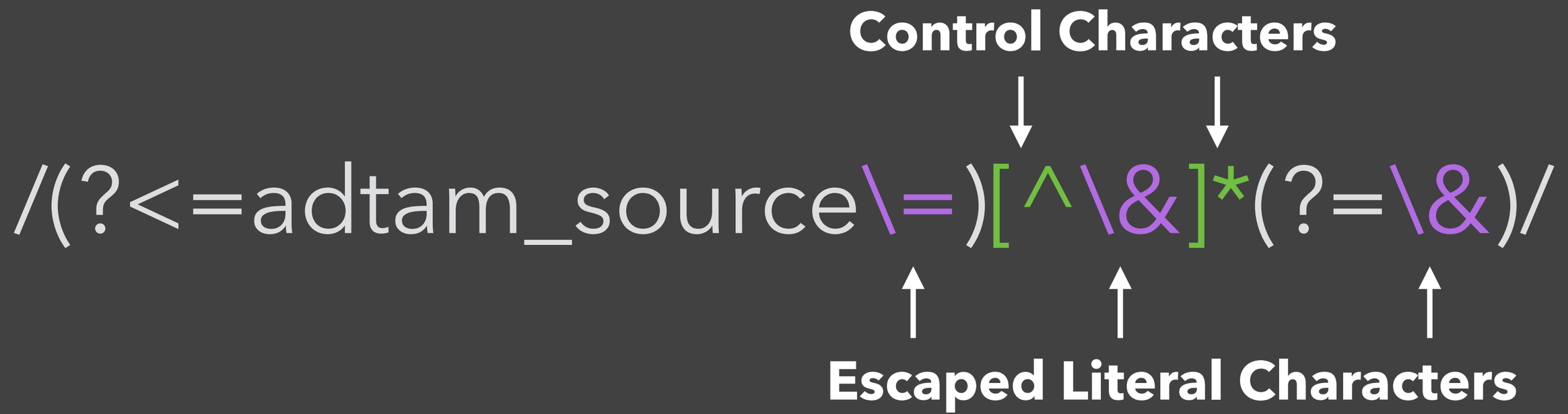
extract *col:* **Screen_Detail** *on:* /(?<=\=)[a-z]+/ *limit:* 2



	Source	Preview	
	abc Screen_Detail	abc Screen...	abc De
	6 Categories	2 Categories	8 Catego
31	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	Nokia
32	adtam_name=holidaypromo1&adtam_source=dynamic&adtam_size=300x250	dynamic	Nokia
33	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	samsung
34	adtam_name=holidaypromo2&adtam_source=mobile&adtam_size=240x400	mobile	Nokia

/(?<=adtam_source\=)[^\&]*(?=\&)/


What (not) to match
/(?<=adtam_source\=)[^\&]*(?=\&)/
Look-behind
Look-ahead


Control Characters
/(?<=adtam_source\=)[^\&]*(?=\&)/
Escaped Literal Characters

Write **once**, read **never.**

after: 'adtam_source=' *before*: '&'

SUGGESTED TRANSFORMS

extract *col*: **Screen_Detail** *after*: `adtam_source=` *before*: `&`

extract *col*: **Screen_Detail** *limit*: 2 *after*: `=` *before*: `&`

extract *col*: **Screen_Detail** *on*: `{lower}+` *limit*: 2 *after*: `=`

Source		Preview	
abc	Screen_Detail	abc Screen...	abc De
	6 Categories	2 Categories	8 Catego
31	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	Nokia
32	adtam_name=holidaypromo1&adtam_source=dynamic&adtam_size=300x250	dynamic	Nokia
33	adtam_name=utarget1&adtam_source=dynamic&adtam_size=180x150	dynamic	samsung
34	adtam_name=holidaypromo2&adtam_source=mobile&adtam_size=240x400	mobile	Nokia

Text Selection: A Language within a Language

Transforms	**Parameter Types**
Split	***Text Selection***
Extract	Column Selection
Filter	Row Selection
Derive	Formula
Header	Enumeration
Pivot	Number
Aggregate	String
Join	Boolean
Union	
…	

Text Selection: A Language within a Language

Transforms	Parameter Types	Text Selection Prepositions
Split	***Text Selection***	on
Extract	Column Selection	from / to
Filter	Row Selection	after / before
Derive	Formula	
Header	Enumeration	
Pivot	Number	
Aggregate	String	
Join	Boolean	
Union		
…		

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

1. User Selects Text(s)

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

1. User Selects Text(s)
2. Tokenize / Generalize

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

1. User Selects Text(s)
2. Tokenize / Generalize
3. Generate Clauses

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

1. User Selects Text(s)
2. Tokenize / Generalize
3. Generate Clauses
4. Combine Clauses

Text Selection: A Language within a Language

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Text Selection Prepositions

on
from / to
after / before

Inference Procedure

1. User Selects Text(s)
2. Tokenize / Generalize
3. Generate Clauses
4. Combine Clauses
5. Filter & Rank Patterns

Source	Preview
ABC CAND_NAME	ABC CAND_NAME2
4,760 Categories	1,349 Categories
COX, JOHN R.	JOHN
ROBY, MARTHA	MARTHA
JOHN, ROBERT E JR	ROBERT
CRAMER, ROBERT E "BUD" JR	ROBERT
BROOKS, MO	MO
COOKE, STANLEY KYLE	STANLEY
SEWELL, TERRI A.	TERRI
HILLIARD, EARL FREDERICK JR	EARL
CHAMBERLAIN, DON	DON
CRAWFORD, ERIC ALAN RICK	ERIC
GREGORY, JAMES CHRISTOPHER	JAMES
CAUSEY, CHAD	CHAD
SMITH, PRINCELLA D	PRINCELLA
GRIFFIN, JOHN TIMOTHY	JOHN
ELLIOTT, JOYCE ANN	JOYCE
SKOCH, BERNARD KURT 'BERNIE'	BERNARD
WHITAKER, DAVID JEFFREY	DAVID
WOMACK, STEVE	STEVE
FALEOMAVAEGA, ENI	ENI

vs.

Source	Preview
ABC CAND_NAME	ABC CAND_NAME2
4,760 Categories	1 Category
COX, JOHN R.	JOHN
ROBY, MARTHA	
JOHN, ROBERT E JR	JOHN
CRAMER, ROBERT E "BUD" JR	
BROOKS, MO	
COOKE, STANLEY KYLE	
SEWELL, TERRI A.	
HILLIARD, EARL FREDERICK JR	
CHAMBERLAIN, DON	
CRAWFORD, ERIC ALAN RICK	
GREGORY, JAMES CHRISTOPHER	
CAUSEY, CHAD	
SMITH, PRINCELLA D	
GRIFFIN, JOHN TIMOTHY	JOHN
ELLIOTT, JOYCE ANN	
SKOCH, BERNARD KURT 'BERNIE'	
WHITAKER, DAVID JEFFREY	
WOMACK, STEVE	
FALEOMAVAEGA, ENI	

Wrangle Language Building Blocks

Transforms	**Parameter Types**
Split	***Text Selection***
Extract	Column Selection
Filter	Row Selection
Derive	Formula
Header	Enumeration
Pivot	Number
Aggregate	String
Join	Boolean
Union	
…	

Wrangle Language Building Blocks

Transforms	Parameter Types
Split	Text Selection
Extract	Column Selection
Filter	Row Selection
Derive	Formula
Header	Enumeration
Pivot	Number
Aggregate	String
Join	Boolean
Union	
…	

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

1. User Makes Selection(s)

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

1. User Makes Selection(s)
2. Infer Parameter Sets

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

1. User Makes Selection(s)
2. Infer Parameter Sets
3. Generate Compatible Transforms

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

1. User Makes Selection(s)
2. Infer Parameter Sets
3. Generate Compatible Transforms
4. Rank & Cluster Transforms

Wrangle Language Building Blocks

Transforms

Split
Extract
Filter
Derive
Header
Pivot
Aggregate
Join
Union
…

Parameter Types

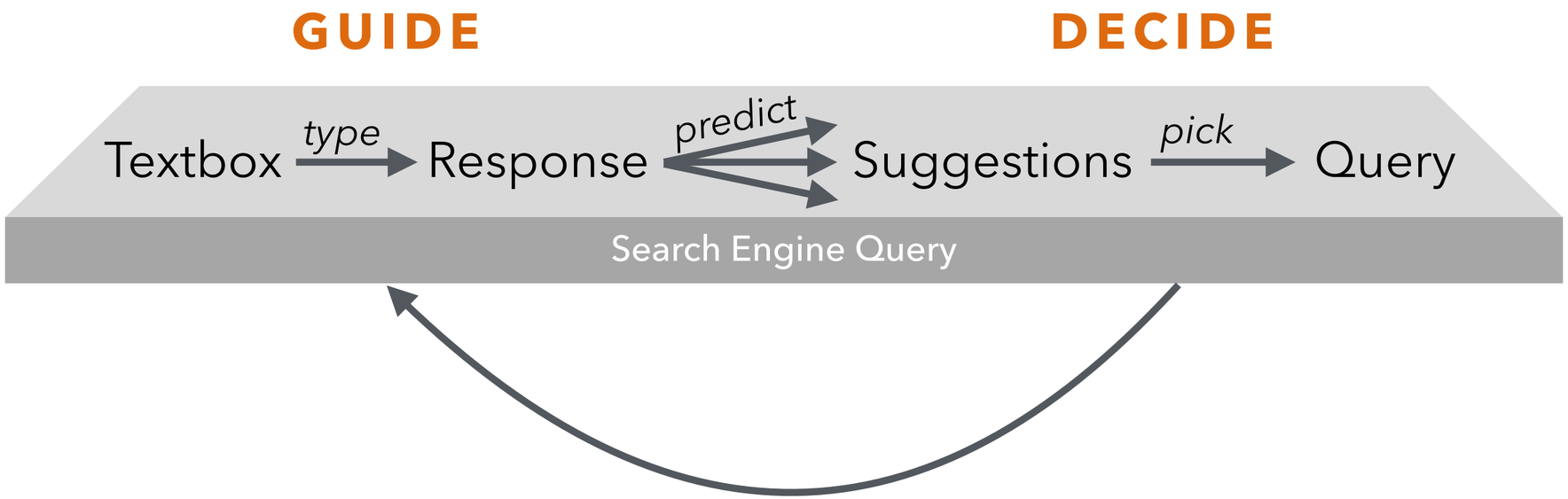
Text Selection
Column Selection
Row Selection
Formula
Enumeration
Number
String
Boolean

Inference Procedure

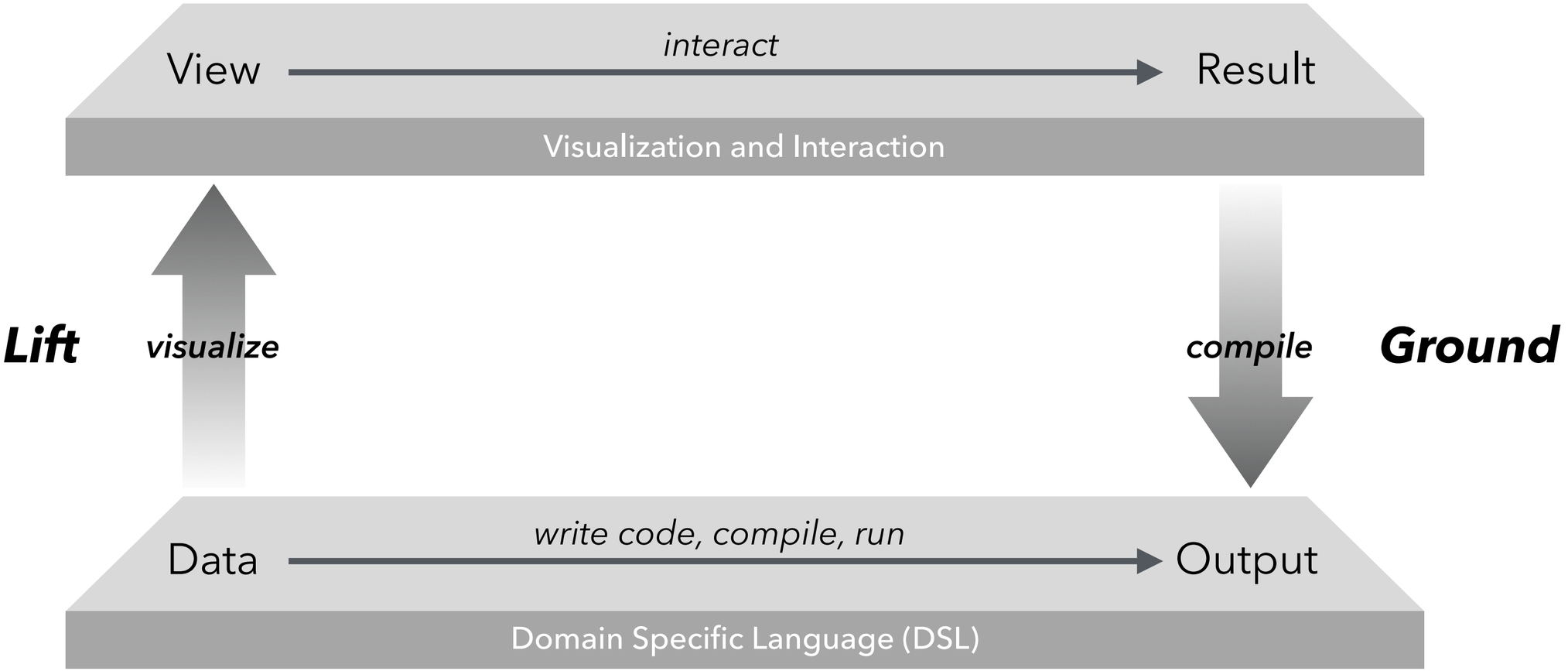
1. User Makes Selection(s)
2. Infer Parameter Sets
3. Generate Compatible Transforms
4. Rank & Cluster Transforms
5. Present Top Results

Predictive Interaction

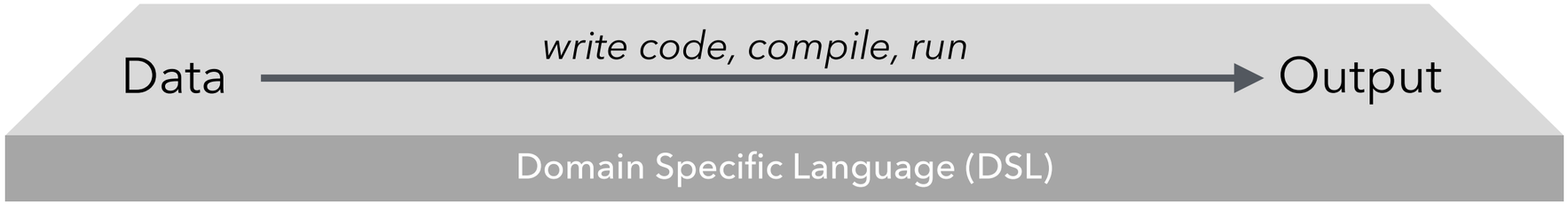
Auto-Complete



Lifting from DSL to Visual Language

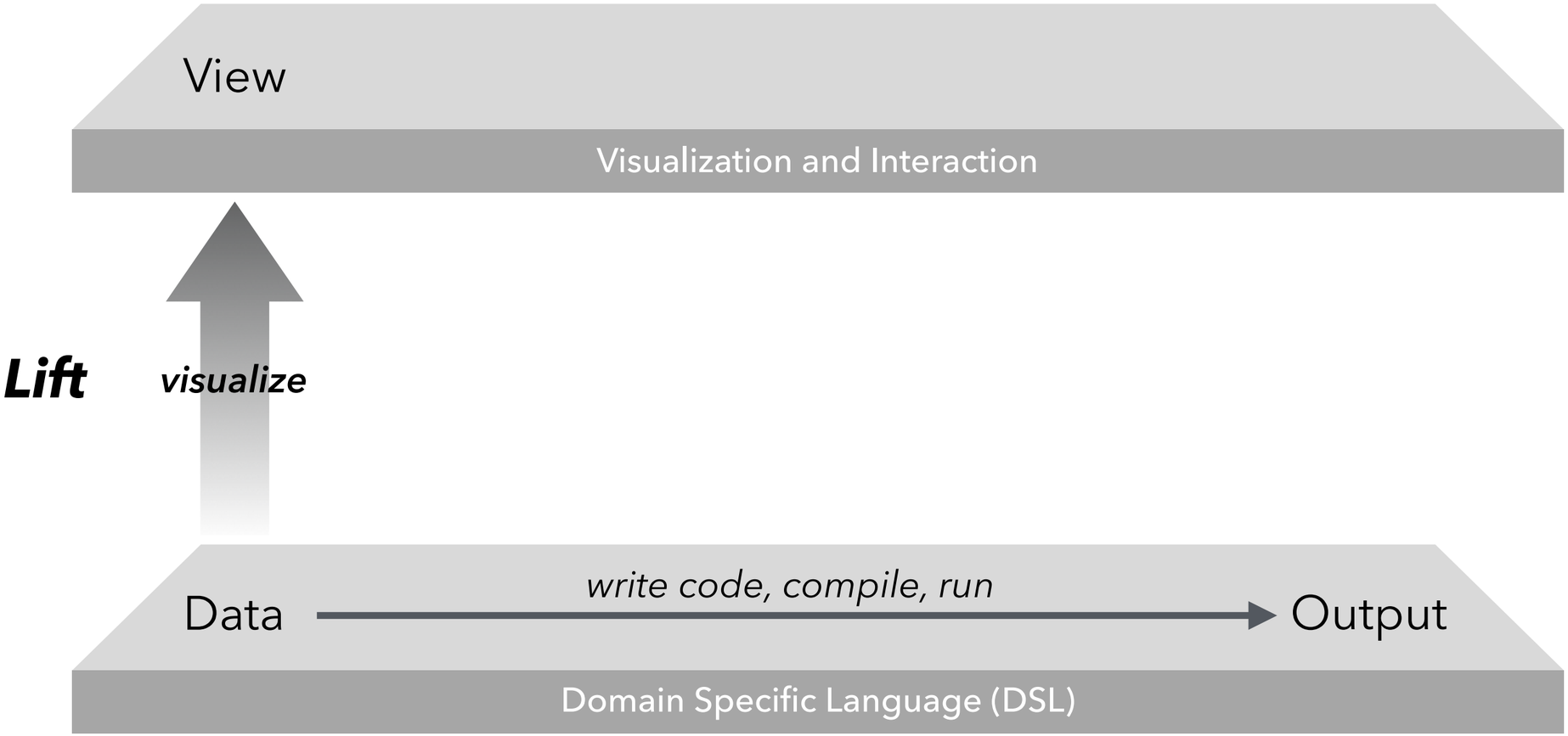


Predictive Interaction

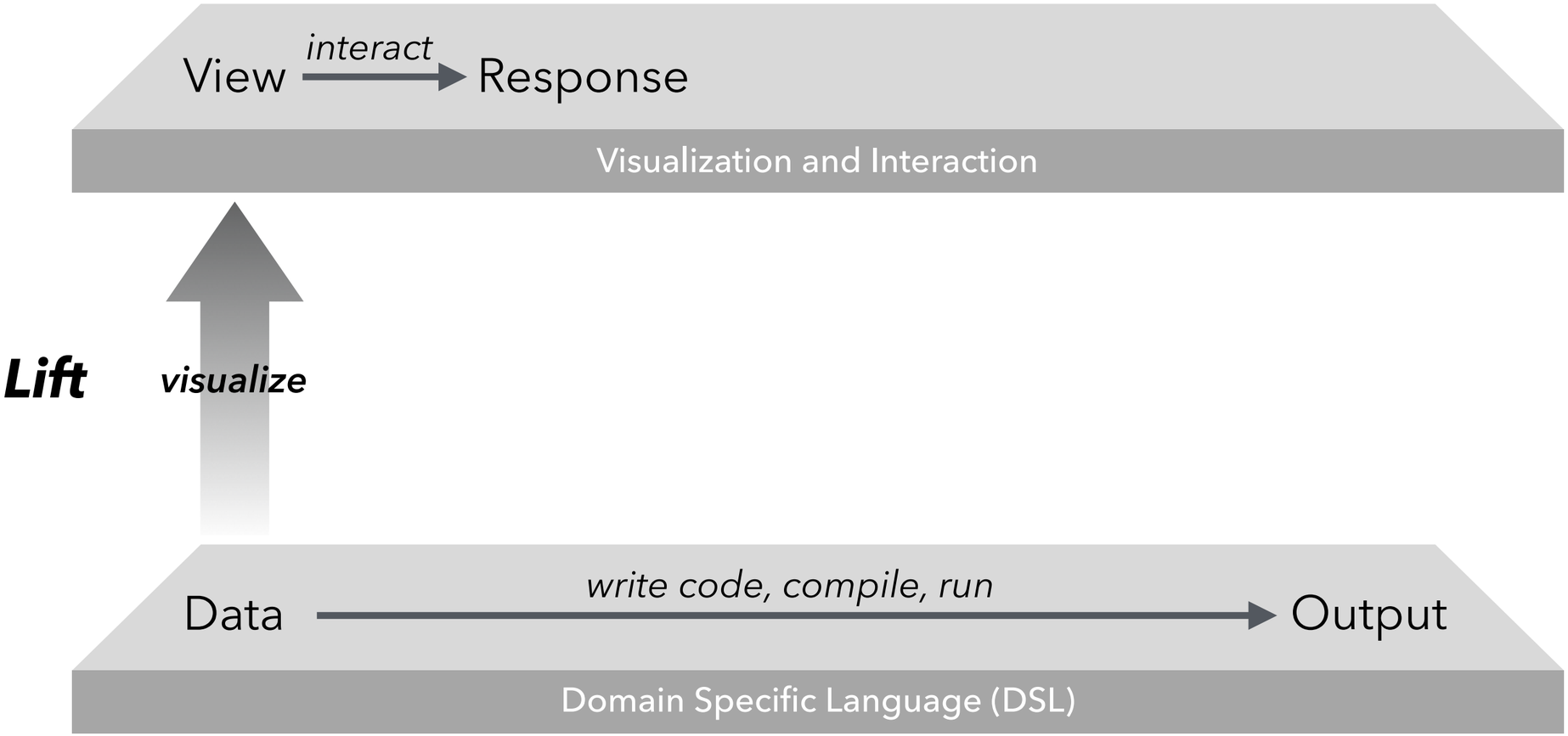


Predictive Interaction

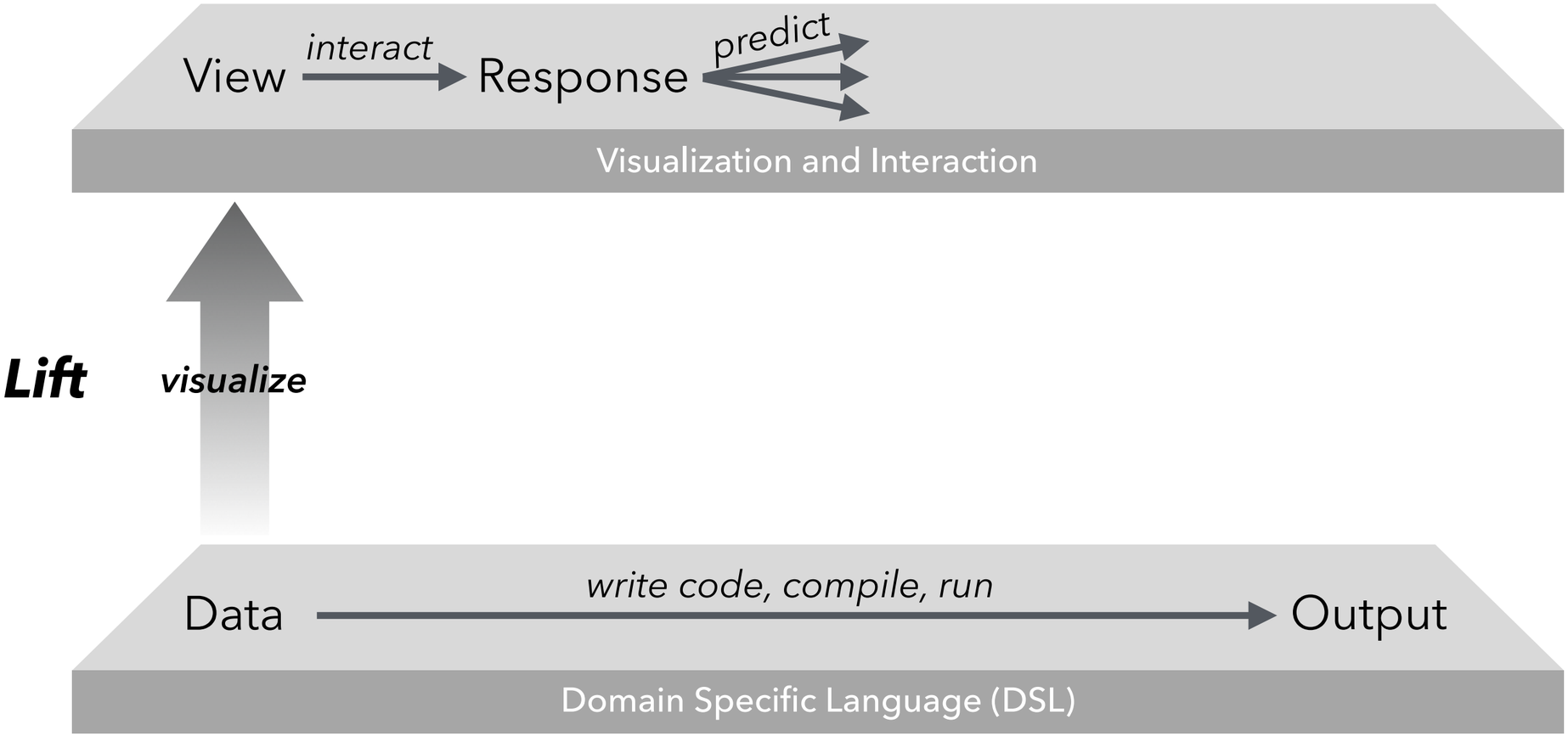
Predictive Interaction



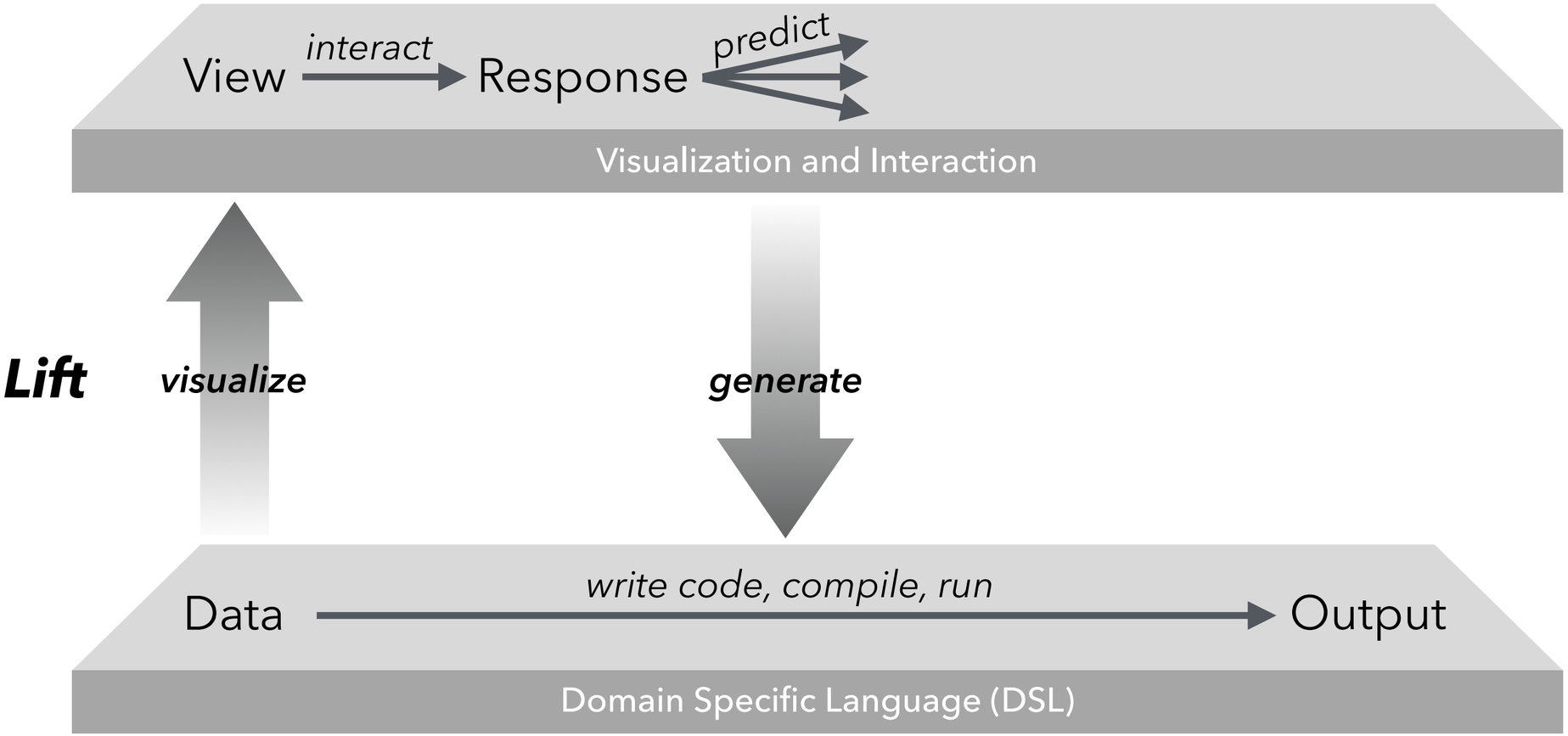
Predictive Interaction



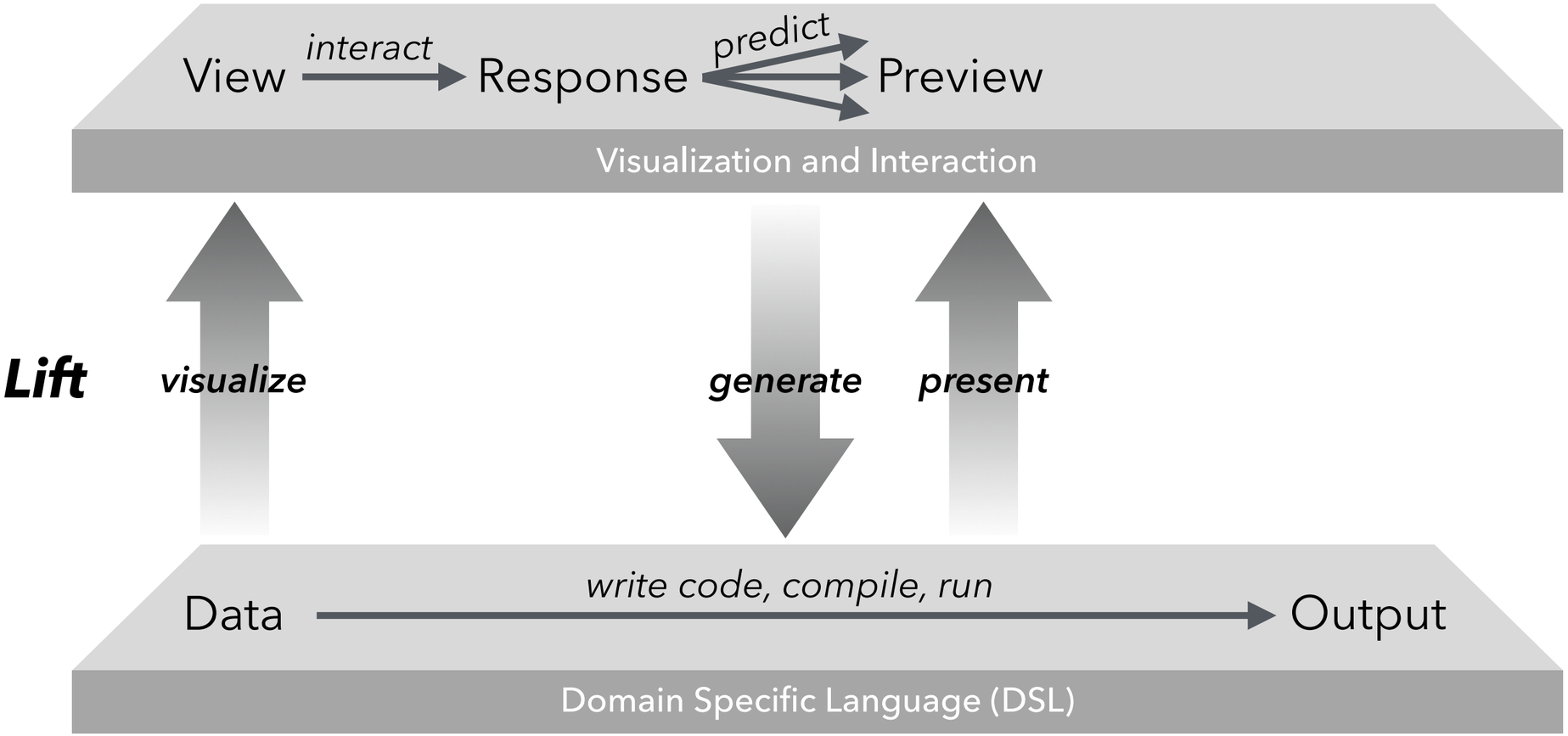
Predictive Interaction



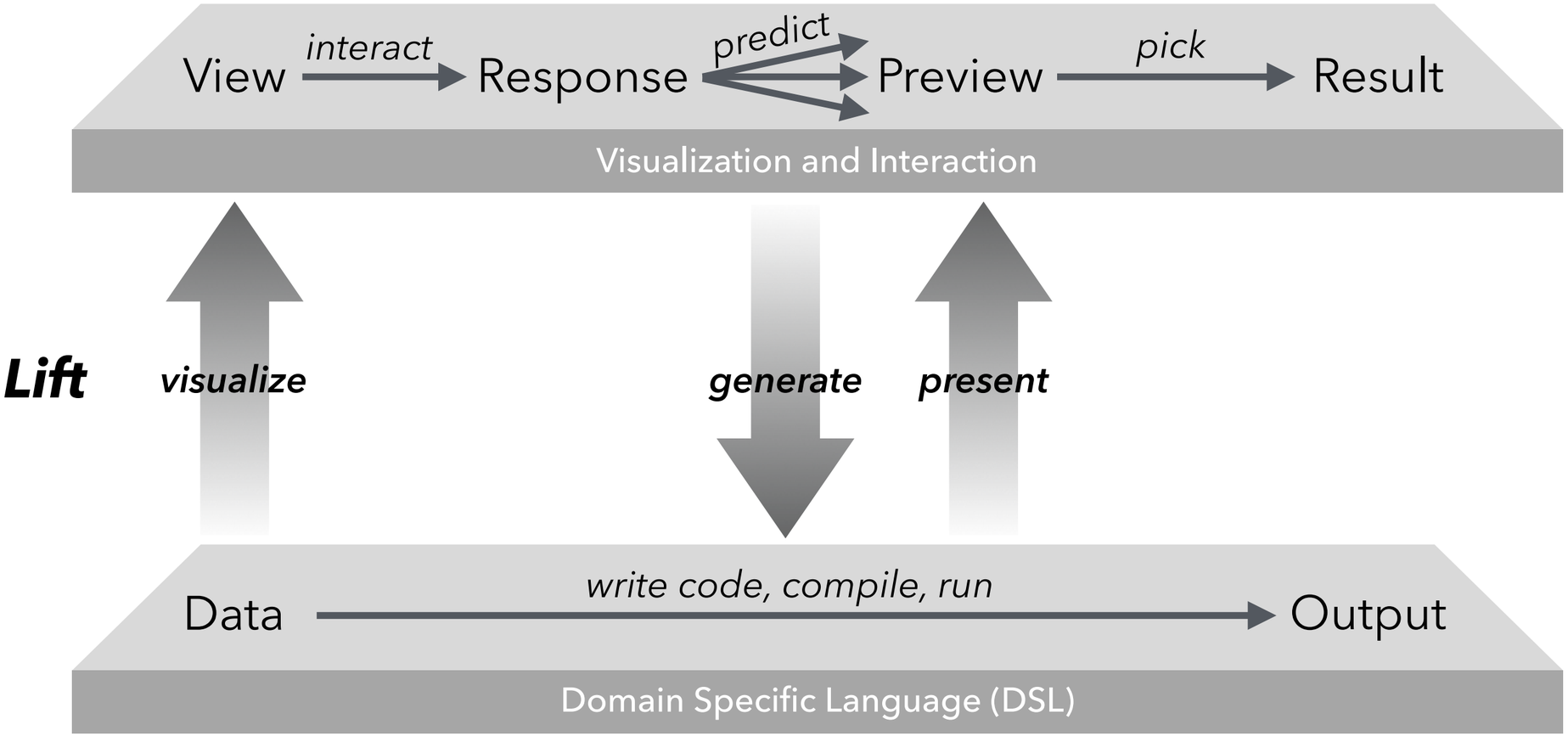
Predictive Interaction



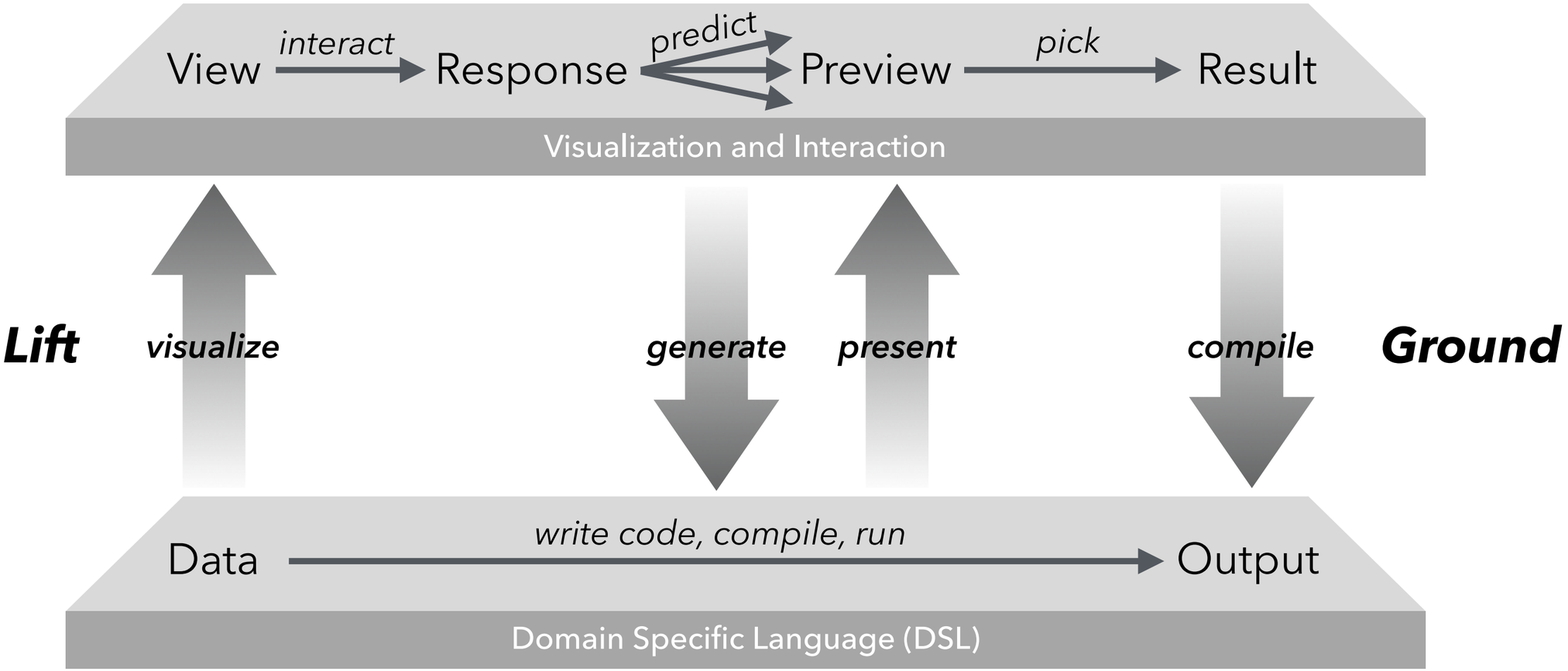
Predictive Interaction



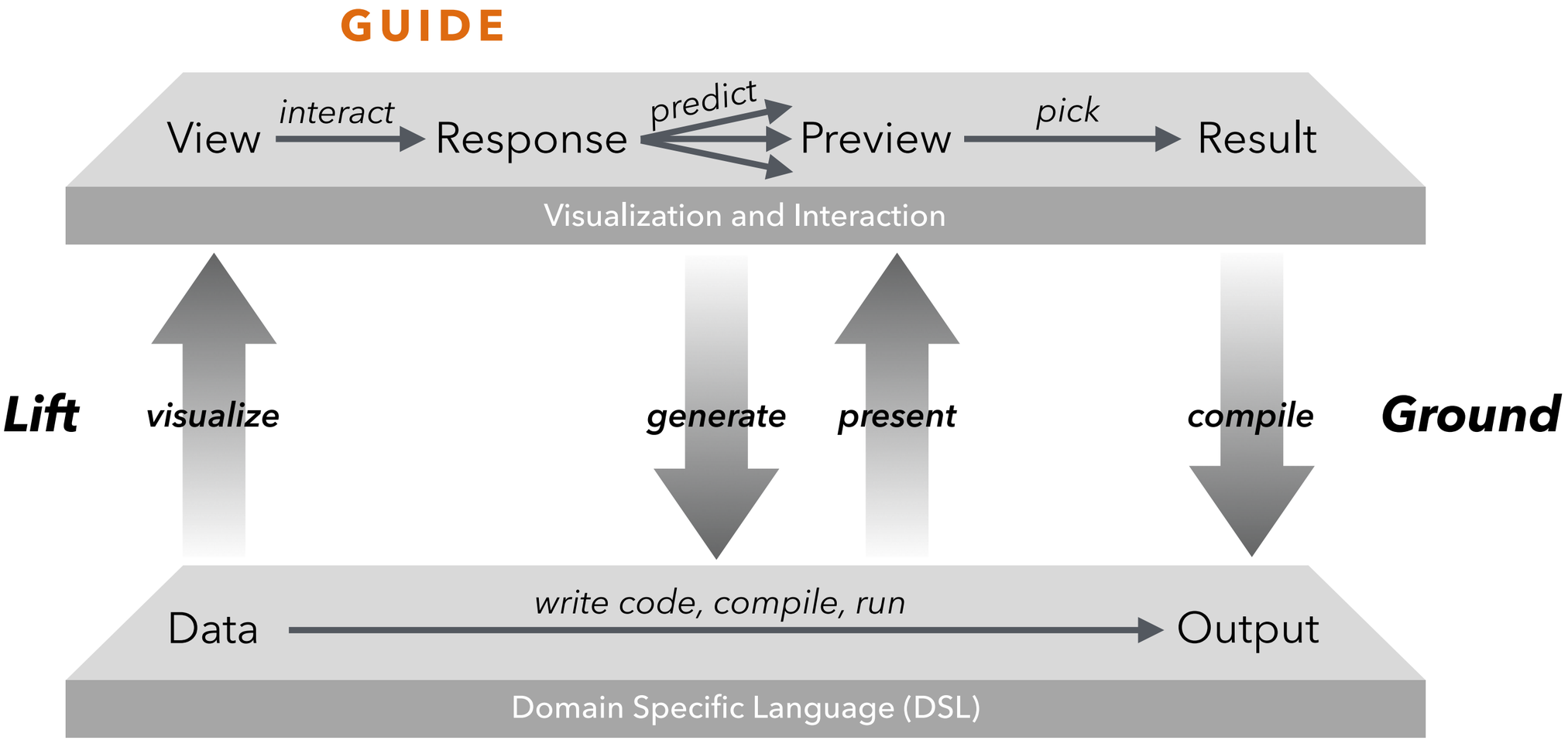
Predictive Interaction



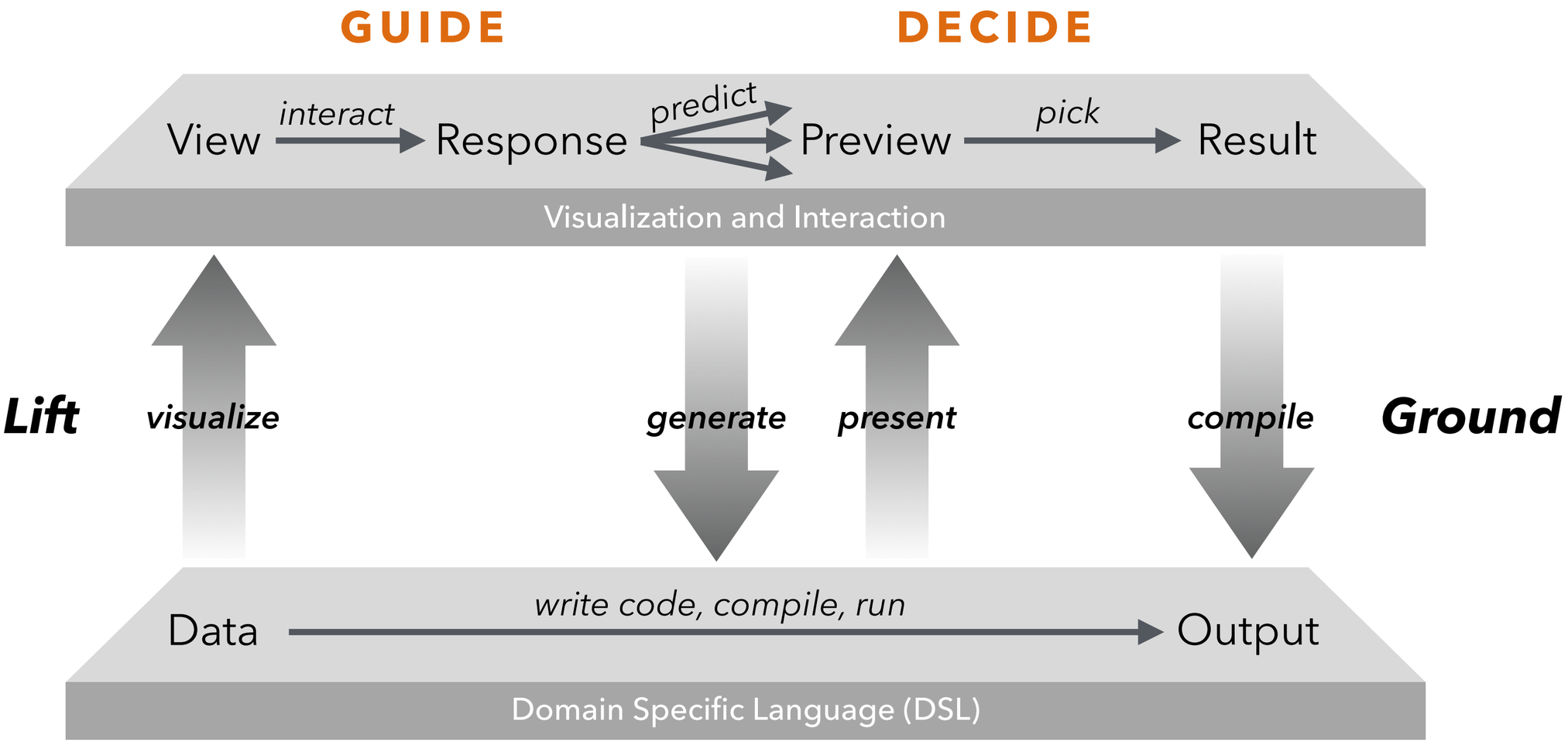
Predictive Interaction



Predictive Interaction



Predictive Interaction



Why Domain-Specific Languages?

Model the task (often as a sequence).

Formalism for reasoning about actions.

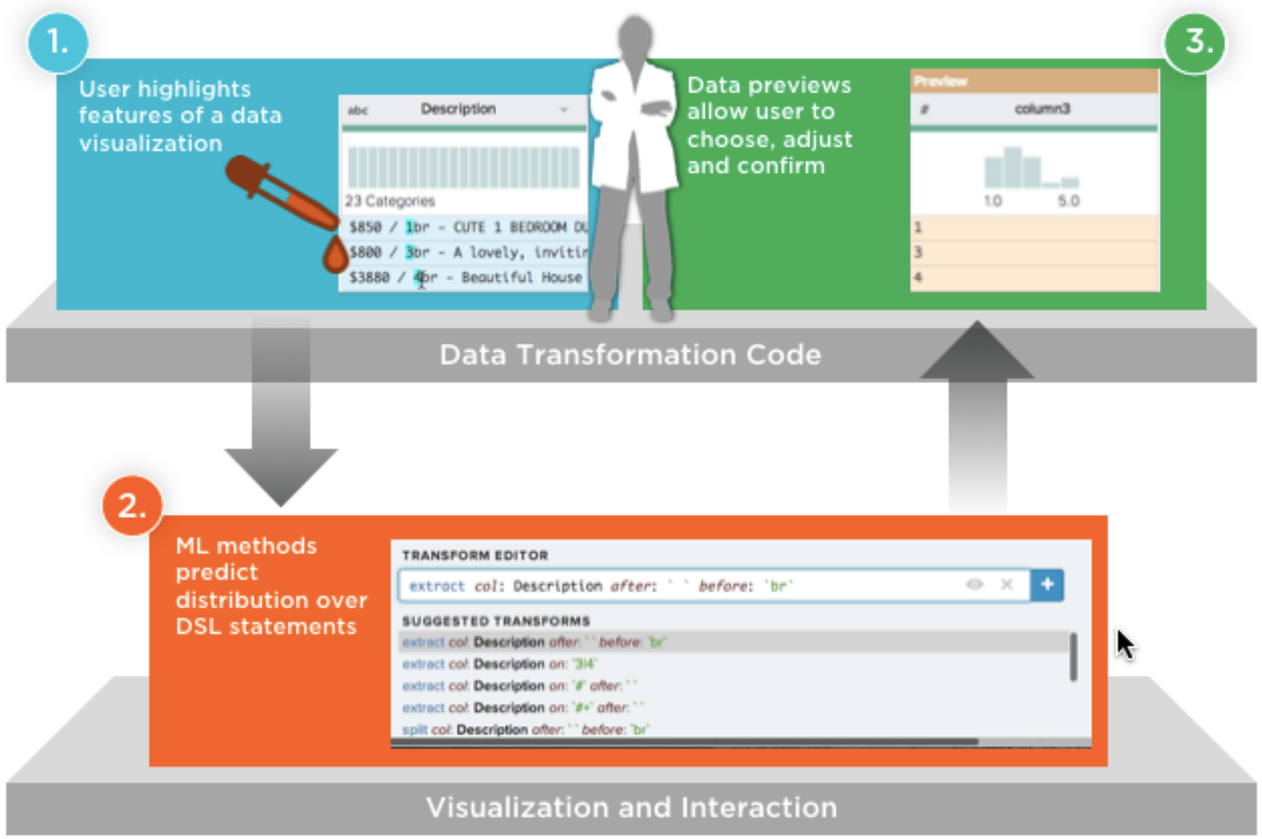
Provides means of learning from usage.

Can be re-applied to new inputs.

Cross-compile to different runtimes.

Necessary Components:

1. Content Representations
2. Language Model
3. Preview Mechanisms



Language Design Considerations

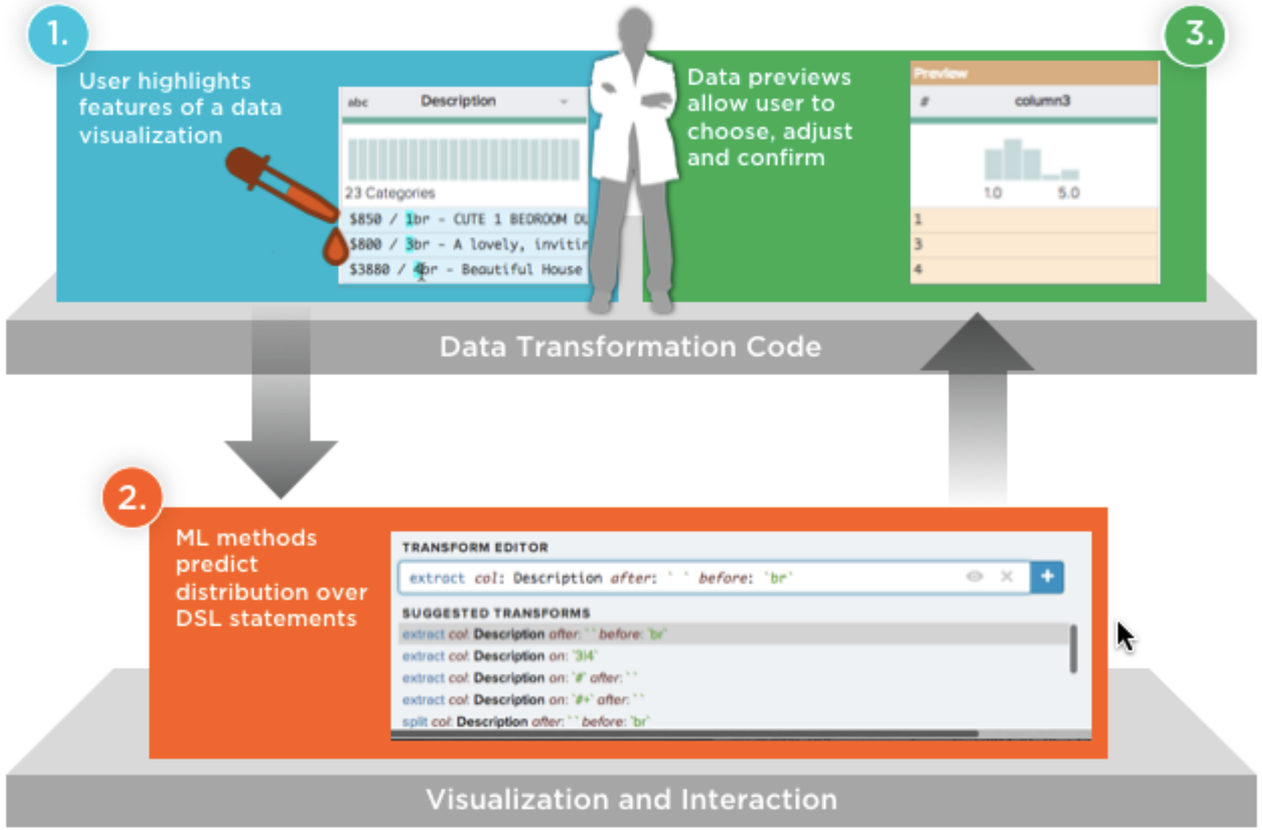
Expressivity. Supports the tasks.

Problem domain fit. Nouns and verbs match domain understanding.

Small surface area. Permits tractable inference, less for users to learn.

Bootstrap ranking. Can the language model provide useful suggestions without extensive training data?

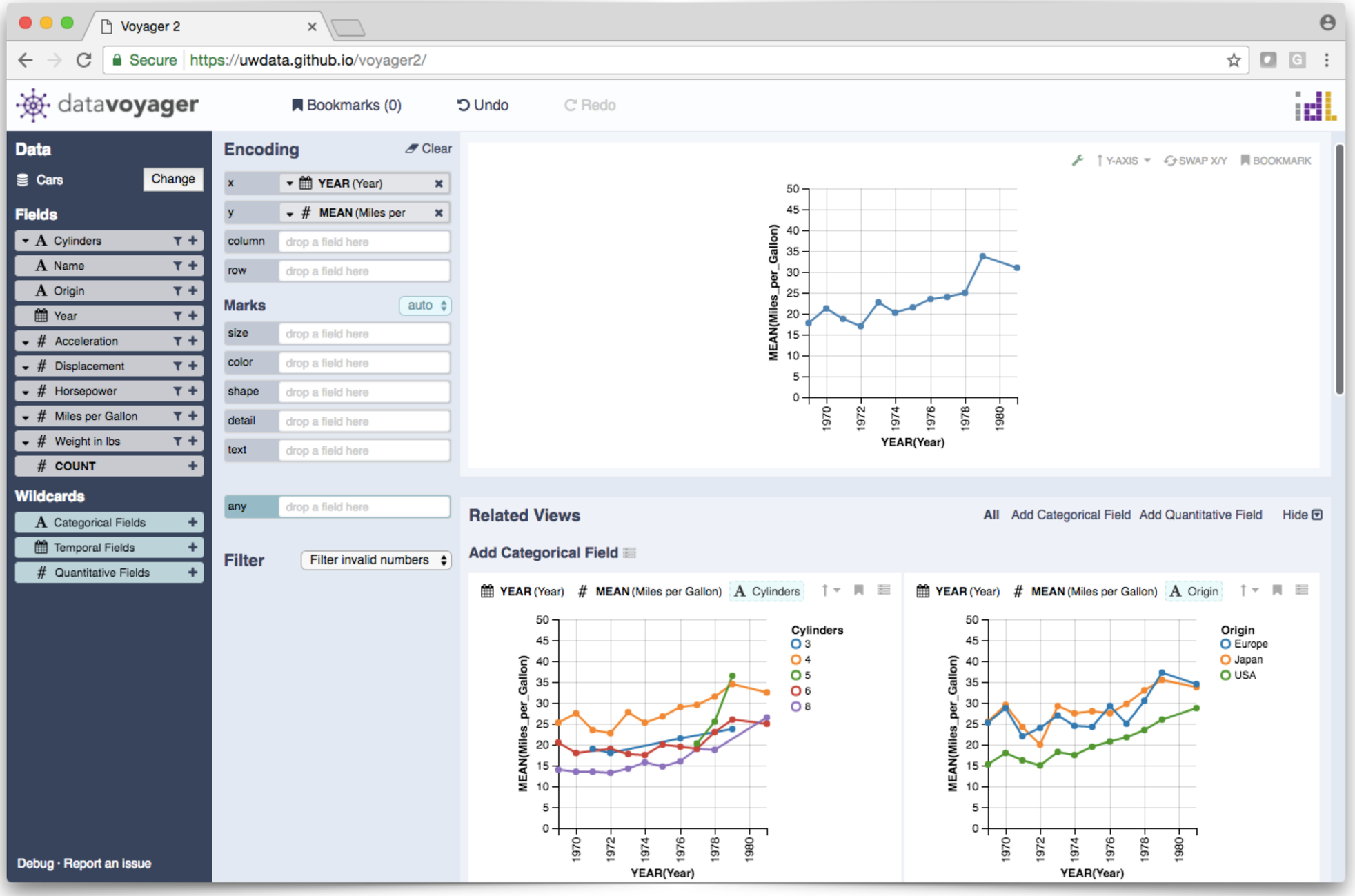
Interfaces that learn. Improving prediction by observing user behavior.



Research Directions

Opportunities & Challenges

Applicable Problem Domains & Language Designs



Voyager. Kanit Wongsuphasawat, Dominik Moritz et al. *InfoVis'15, CHI'17*

Opportunities & Challenges

Applicable Problem Domains & Language Designs

Mixed-Initiative Interaction, Proactive Suggestions

User Performance Cliffs, Error Handling

Learning from Usage; How to Interpret Observed Input?

Reusable Abstractions for UI, Language & Inference?

Predictive Interaction

Jeffrey Heer @jeffrey_heer
U. Washington / Trifacta