TABLEAU CONFERENCE
Turning Questions into Answers with Tableau

Ross Bunker
Principal Engineer
Tableau Software
Helping People See and Understand Their Data

Tableau’s Mission
"Understand: Perceive the significance, explanation, or cause of"

Oxford English Dictionary
Overview

Anatomy of a Question

Answering Questions using Tableau

Summary and Future
Anatomy of a Question
Anatomy of a Question

Warning! Grammar ahead....
Anatomy of a Question

‘Things’ the question is about
“What were the bestselling cars of 1990?”

‘Attributes’ of ‘things’
“What types of cars sold best in 1990?”

Sometimes a placeholder for the verb
“What happened to car sales in 1990?”

Also ‘which’
“Which cars sold best in 1990?”
Anatomy of a Question

WHO

Semantically a person
In terms of data, just another ‘What’
(But semantics are definitely interesting)
Anatomy of a Question

Time-series => use a timeline

When did things happen?
“When did sales increase?”

Prediction
“When will sales level off?”
Requires inference, i.e. forecasting or modelling

Time-series is not always indicated by when
“In what year did the most cars sell?”
Anatomy of a Question

Location analysis
“Where are my sales the strongest?”

Not always a map
“Where do I sell the most staplers?”

Geo-spatial needed for locality questions
“Does proximity to California affect delivery times?”

Spatial is not always indicated by where
(as above)
“What states had the highest sales?”
Anatomy of a Question

Causation inference
This what people are for!
aka Job Security
Anatomy of a Question

Causal versions are like ‘Why’
“How come my profits went down?”

Quantity, magnitude, frequency
“How many staplers did we sell last year?”
“How far do my customers live from my stores?”
“How long between ordering and delivery?”
“How often do my stores run out of supplies?”

Focus shifts to the rest of the question
“sell”, “live”, “between”, “run out”
Anatomy of a Question

Introductory words aren’t much help

Same question can be asked with different introductory words

“Where were car sales the highest?”

“What states had the highest car sales?”

Focus on nouns, verbs and clauses
Anatomy of a Question

Often defines or points to the measurement Action to measure
“How many cars were purchased last month?”

Or measure is the subject/object of the verb
“How did sales do this past year?”

Comparison
“How does marketing spend affect sales?”
“How did this month’s sales compare to last month’s?”
“Which products sold better/worse this month?”

Change
“How many salaries went up/down this year?”
“When did sales start to rise?”
“How long after purchase do cars lose the most value?”
Anatomy of a Question

Dimensions in your question
“How many cars sold in each state?”

Not just the subject of the question
“Which employees sold the most cars in each city?”

Sometimes implicit with the verb
“What were sales like last year?”
“What were home sales like last year?”
Anatomy of a Question

**Domain**—values from the data

- “Washington”, “Seattle”, “Europe”
- “Stapler”, “Ford”,
- “Large”, “Blue”, “Female”
- “August 27, 1995”, “3:00p”

Often preceded by keywords

- “in” <place>
- “on” <date>
- “among” <choices>

Sometimes qualify the dimension
Anatomy of a Question

Adjectives
“How many customers made their first order last year?”
“How did blue cars sell in the year 1990?”

Superlatives
“What was the bestselling car this year?”
“Which cars sold the most in each state?”

Qualifiers
“That”, “who”, “which”, “in”
“For customers that bought staplers, how many also bought chairs?”
Anatomy of a Question

Relative clauses

Important information about nouns

“Which customers who bought staplers also bought chairs?”

May be used with or without comma

“Of the cars sold last year, how many were bought by someone who also bought a car the year before?”
Anatomy of a Question (cont’d)

Special Comparisons

Change over time
“How did sales this year compare to last year?”

Spatial analysis
“How do employee costs relate to urban areas?”

Part/Whole
“How does each state’s sales compare to the average for all states?”
Answering a Question with Tableau
4 Kids + 2 Adults + 3 Drivers ≠ 2 Cars
When will the family Tesla arrive?
Choosing your data

Step 1: Identify ‘things’ and ‘attributes’
Choosing your data

Step 1: Identify ‘things’ and ‘attributes’

Step 2: What are you measuring?
Choosing your data

**Step 1:** Identify ‘things’ and ‘attributes’

**Step 2:** What are you measuring?

**Step 3:** Identify your data
Interlude: LOD and your DataSource
Interlude:
Target audience
Process

1. Identify the verbs and nouns that indicate what you want to measure
2. Figure out the other nouns in your question and find the fields for them
3. If you have any qualifiers or relative clauses, build filters to reflect them
   Use Sets or LOD calculations for relative clauses in many cases
4. Map analytical style to appropriate Tableau features
   Maps for showing spatial proximity
   Timeline for time based questions
   Sorting for superlatives/comparison
   Totals and Reference Lines for part/whole
   Table Calculations for comparison over time or part/whole
5. Augment an audience targeted dashboard with appropriate features
   Consider filter controls, parameters and actions to allow users to customize the question
# Mapping to Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Measurements</th>
<th>Superlatives</th>
<th>Qualifiers</th>
<th>Relative Clauses</th>
<th>Change over time</th>
<th>Spatial Awareness</th>
<th>Part/Whole</th>
<th>Audience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregation</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sorting</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Filters</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sets/LOD/Context</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Table Calculations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Maps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Totals/Ref Lines</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Parameters, Actions, Filter Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
The Future

Natural Language Processing
- Automatically build vizzes from questions
- Customize vizzes using natural language

Richer Semantic Model
- Beyond Geography: Dollars, URLs, Customer
- Allow customization

Richer Data Model
- Understand context of measurements
- Express more questions with a single datasource
Top of the Table | Winning with table calcs
Thu | 10:45a – 1:15p | L2 – 208

As good as it sets
Wed | 1:45p – 4:15p | L2 – 294
Please complete the session survey from the My Evaluations menu in your TC18 app.
Thank you!

Contact or CTA info goes here