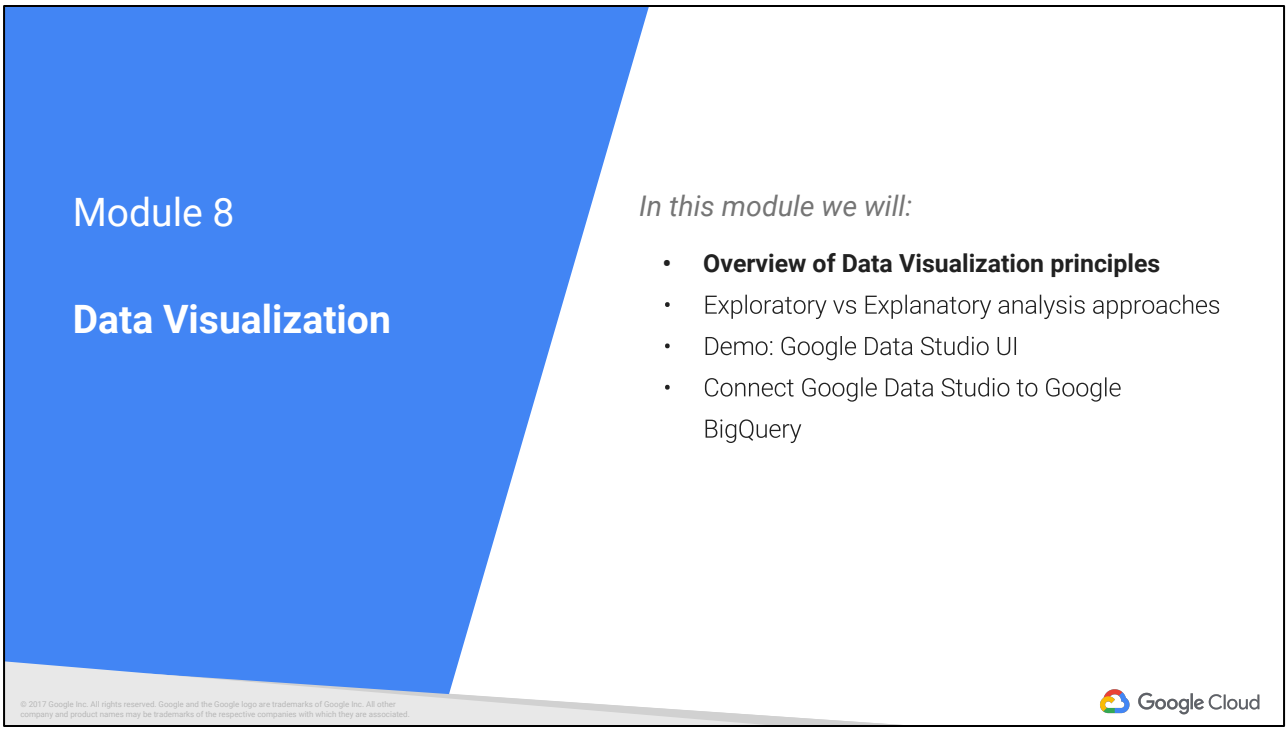
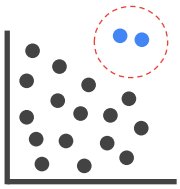


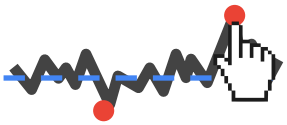
Authors: evanjones@ rachelkatz@ rge@
Owners: evanjones@

One of the key outputs data analysts create are insightful reports for your audience. In this data visualization module we will cover a little visualization theory and best practices and then we'll introduce Google Data Studio as one of the visualization tools in your toolkit for creating actionable reports.

Use visualization to clearly and concisely present insights



Visualizing a dataset allows you to spot hidden trends



Interacting with a dataset visually is often faster than writing SQL

Deliver powerful insights to your audience through reports

Get scalable performance as your dataset grows with BigQuery-backed visualization tools

Google Cloud

Visualization helps make sense of data

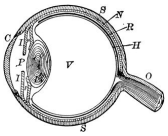
Visualization tools make it easy to “see” data interactively

Visualization tools that support BigQuery can give “small dataset” performance against billion+ row data

Build visualizations without writing SQL

Visualization Theory: Perception

Stimulus



Eye

Brain

Preattentive processing
Attentive processing

Sensation
Physical process

Perception
Cognitive process

Google Cloud

We don’t see things with our eyes, we see things with our brains.
Eyes record light, translate it into electrical signals, and pass it to your brains, where the message is understood.

Preattentive processing:

- Iconic memory: <1s, automatic, unconscious

Attentive processing:

- Short-term memory: 4 chunks of information at most (color, shape, size, number, etc.)
- Long-term memory: holds more information, not our focus today

Visualization Theory: Count the Fives

69750429347493732418605783578
58728294974654487818676453214
24439684634233529867321903875
65878893745390932975659391732
14725920189374476564722175652

Google Cloud

Answer: 16
Took so long because you had to use focused attention system and scan every single row and value

Visualization Theory: Count the **Fives**

6975042934749373241860578357858728294974654487818676453214
2443968463423352986732190387565878893745390932975659391732
1472592018937447656472217565 2

Google Cloud

Much easier with the help of preattentive processing - features extracted quickly, effortlessly, in parallel without any attention focused on it

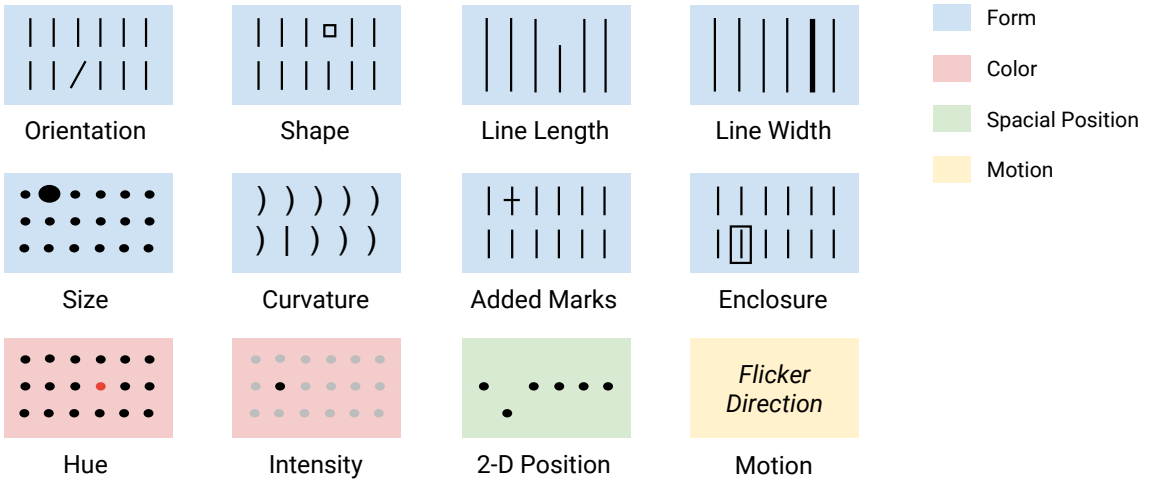
Visualization Theory: Count the **Fives**

697504293474937324186057835785872829497465448781867645321424439684634233529867321903875658788937453909329756595917321472592018937447656472217562

Google Cloud

Even easier with visual encoding

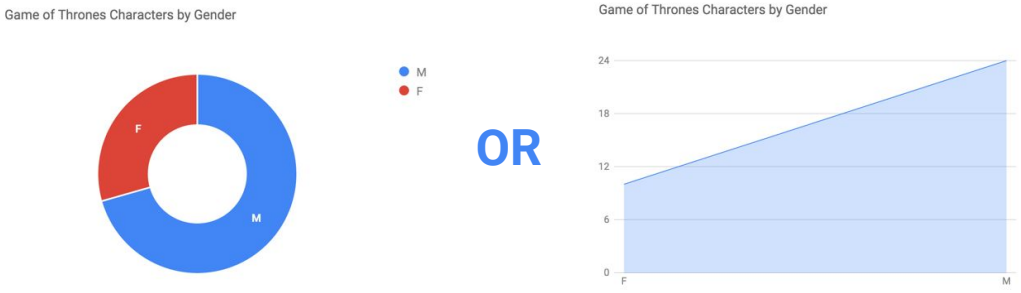
Visualization Theory: Preattentive Attributes



Google Cloud

Use encoding to highlight the relevant information and help your audience understand your message in their preattentive space. That leaves the attentive space for them to listen to what you're saying!

You Choose: Effective or Ineffective (or Wrong) Visuals

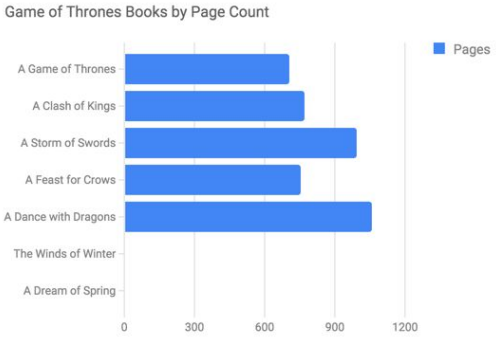


Google Cloud

Donut Chart for quick comparison of book characters by gender is likely the better visual

Line graphs are mainly for time series and does not make sense to use here.

You Choose: Effective or Ineffective Visuals



OR

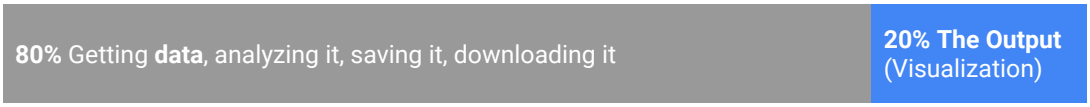
Google Cloud

Horizontal bar chart allows you to have dimensions with valid 0 values (like two books that have not been written yet).

Donut charts can get crowded with multiple dimensions.

The 80/20 Rule

What you spend your time on:



What your audience actually cares about:

Google Cloud

Your audience is likely only to see the end product of your work. Your visuals and reports often carry more visual weight than your SQL scripts and analysis.

Visualization Core Concept: Dimensions and Measures

	Description	Examples
1 **Dimensions**	A field that can be considered an independent variable. Usually contains qualitative, categorical information	• Name • Location • Part Number # • Job Title
2 **Measures**	A field that is a dependent variable; that is, its value is a function of one or more dimensions. e.g. any field containing numeric (quantitative) information	• Revenue • Salary • Expenses • Count of Errors

Class Question

Which of the below are measures?

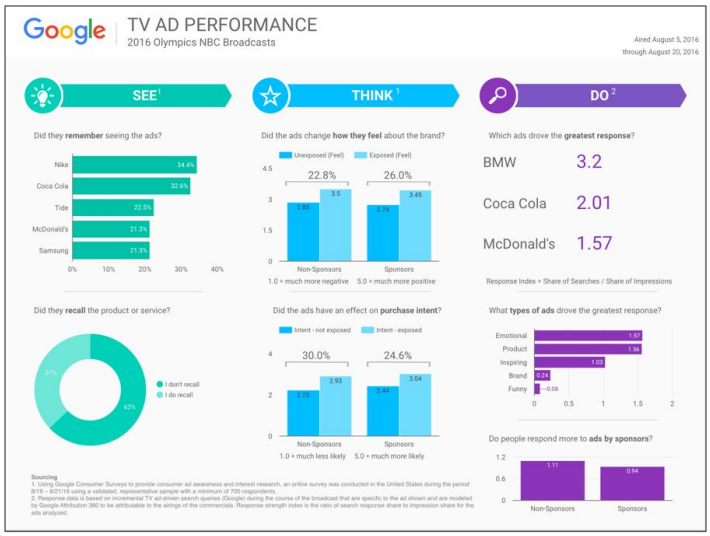
1. Phone Number
2. Employee ID
3. Age
4. Date of Birth
5. Tenure at Work (in years)
6. Job Title

Remember, measures are usually quantitative fields

Google Cloud

Age and Tenure (you can do math on them)... in reality, you can do COUNTs of unique Job Titles and that would become a measure

Reports transform data into information



Tell a clear story with your data

Share and collaborate on reports with others

Google Cloud

Here's an example of an actual Data Studio report. You don't have to understand the numbers, but just on the surface, what is this report trying to do?

Role of the report:

- Visualize data
- Restrict data
- Share with viewers and editors

Module 8

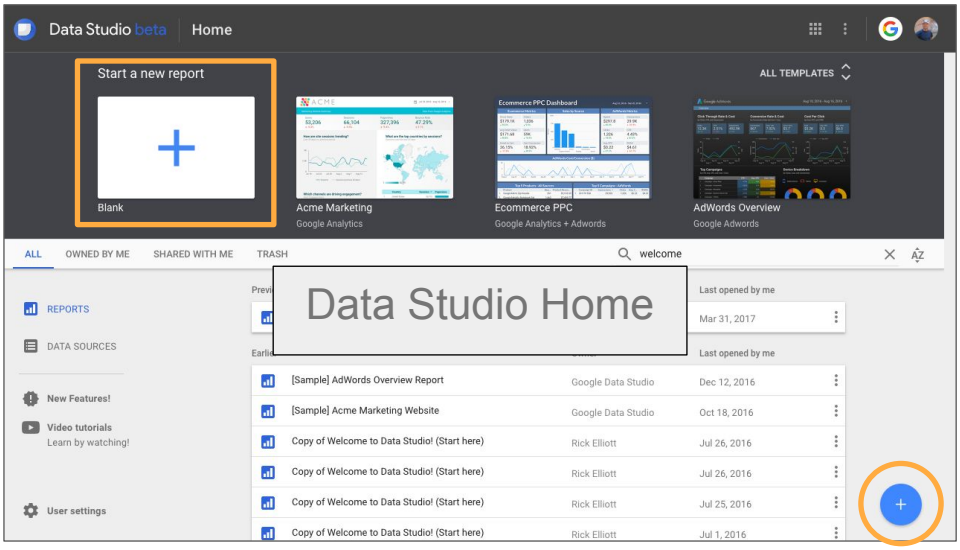
Data Visualization

In this module we will:

- Overview of Data Visualization principles
- Exploratory vs Explanatory analysis approaches
- **Demo: Google Data Studio UI**
- **Connect Google Data Studio to Google BigQuery**

Google Cloud

Create new reports in the Data Studio UI



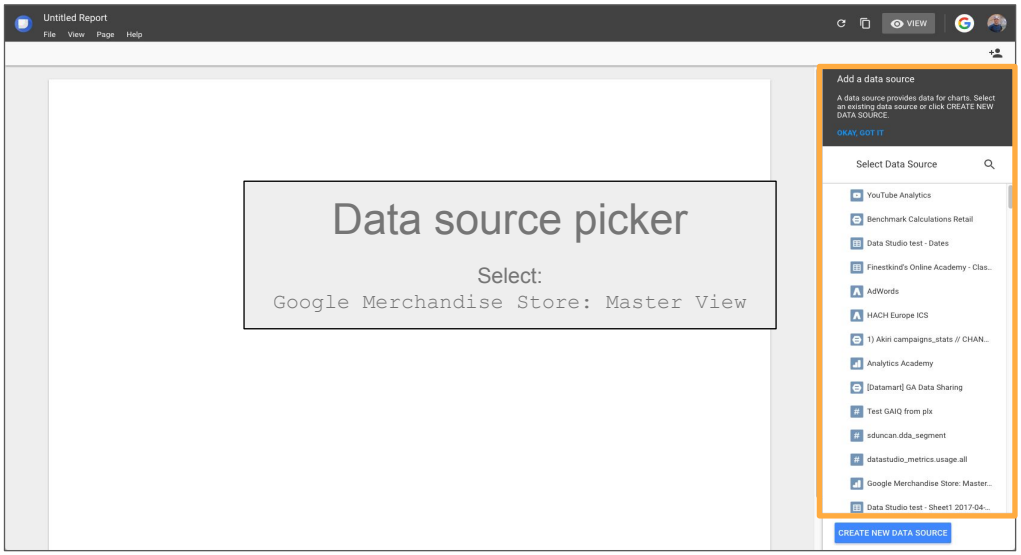
Google Cloud

This is the Data Studio Home page. We'll come back for a quick tour of some of the features in a bit.

For now, there 2 ways to create a new report from scratch: from the templates panel on top

Or from the button in the lower right.

Select data sources to build your visualizations



Google Cloud

The first thing you need to do is tell Data Studio where your data is coming from. That is known as the "data source"

A Data Studio report can have any number of data sources, but we'll just start with one.

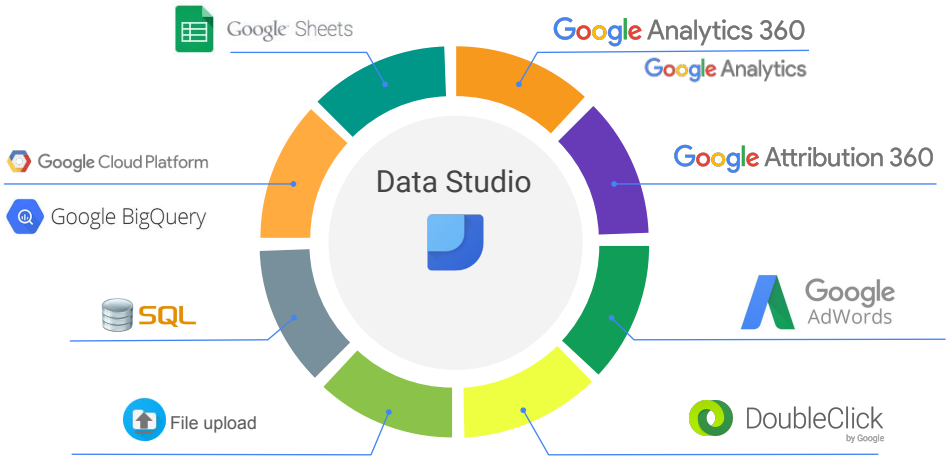
The data source picker shows all the data sources you have access to.

Browse or search for *Google Merchandise Store: Data Studio 101*

I've shared this data source with you.

Select *Google Merchandise Store: Data Studio 101*

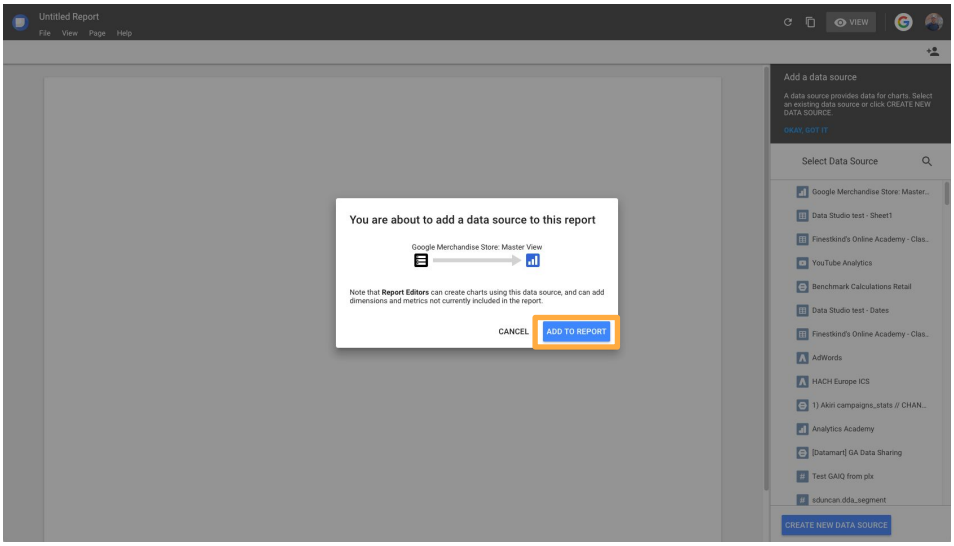
Connect to multiple different types of data sources



Google Cloud

Note that you can have any or all of these data sources in a single Data Studio report

Add the data source to your report



Google Cloud

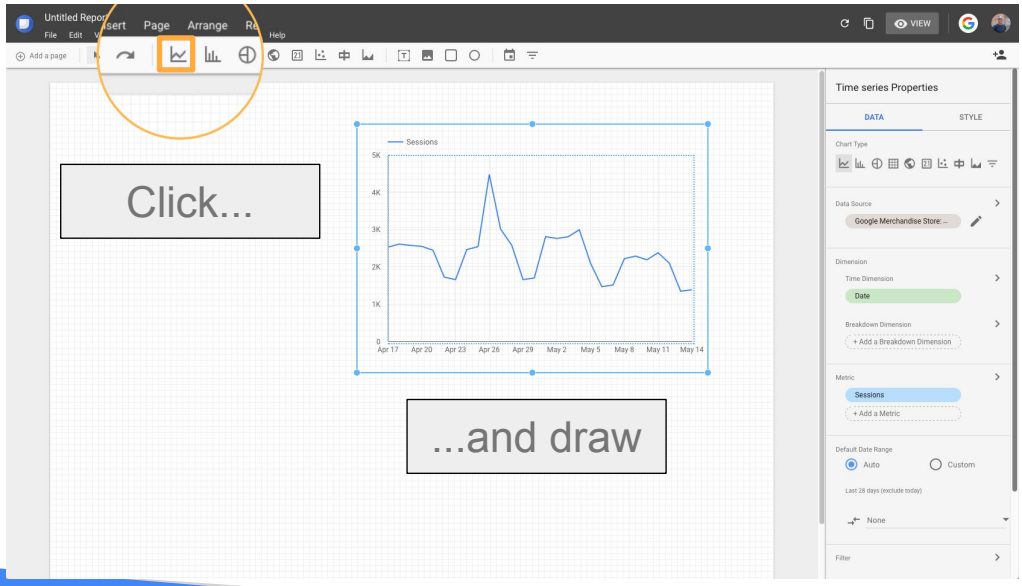
Since Data Studio reports can be shared, you should be aware of the ramifications of adding a data source.

When you add a data source to a report, other people who can view the report can potentially see all the data in that data source. And anyone who can edit the report can use all the fields from any added data sources to create new charts with them.

We'll talk about how to control access to data and sharing later on in this course.

Click ADD TO REPORT

Create charts to visualize data relationships



Google Cloud

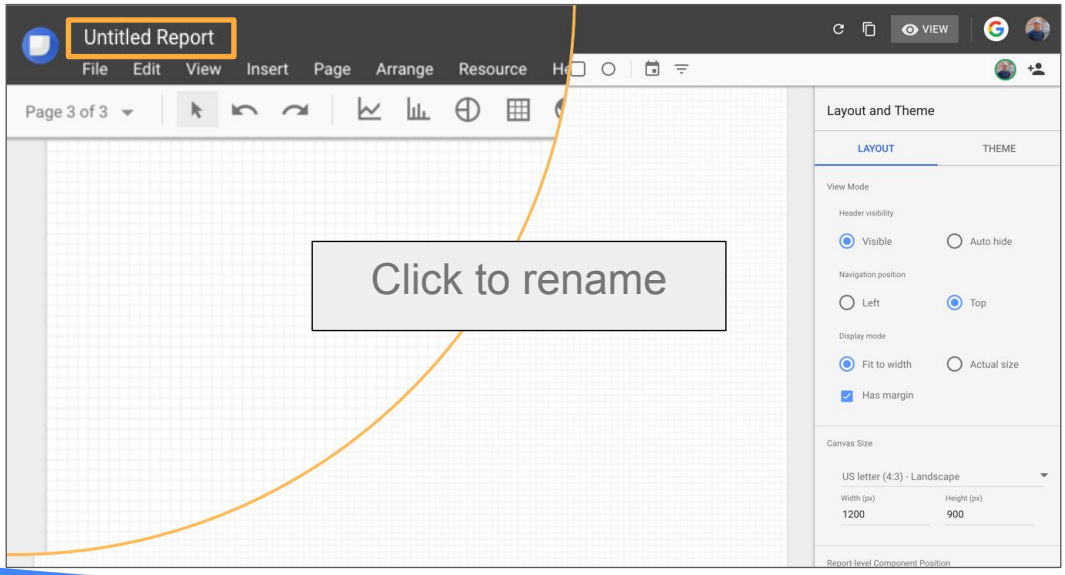
Data Studio is "click and draw", not "drag and drop".

Select the chart type, then draw a box where you want it to go.

Data Studio automatically adds some appropriate fields to the chart. The chart properties panel on the right is where you configure the chart. We'll come back to that in a bit as well.

Add a time series to the report

Add a descriptive name to your report



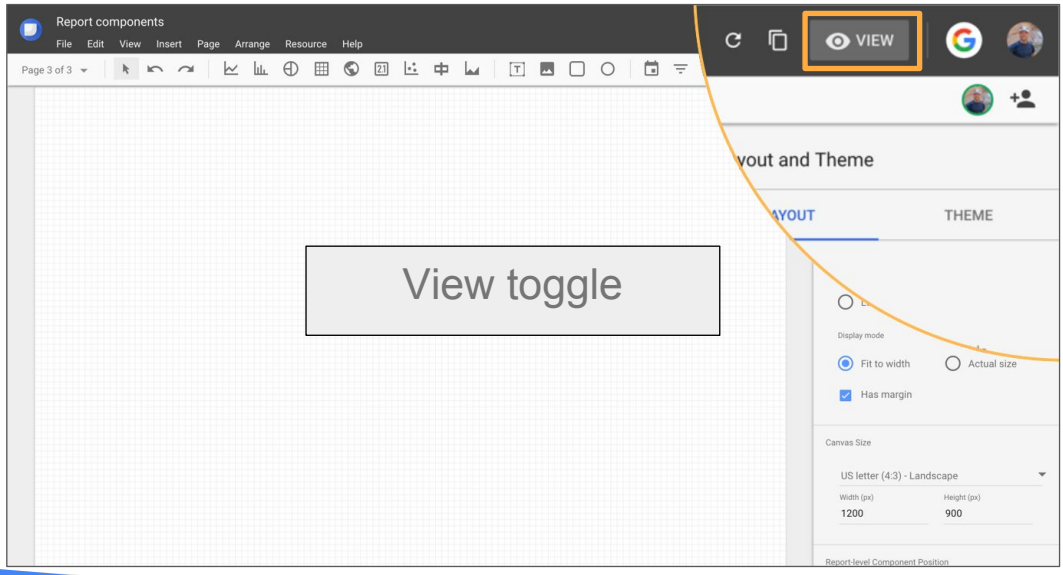
Google Cloud

Give your report a name.

One thing to watch out for, since Data Studio is based on Drive, you can have duplicate file names.

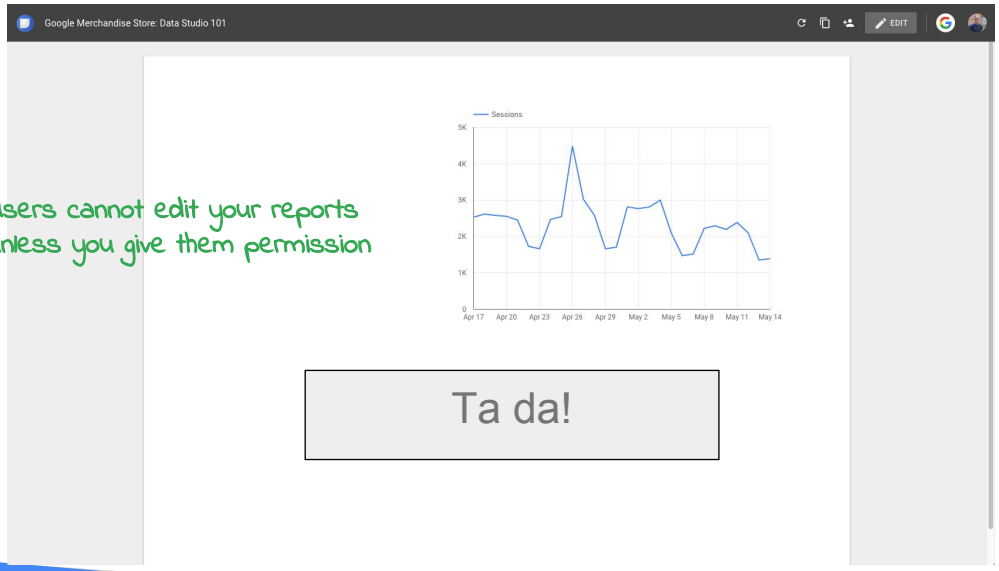
Name the report *Data Studio 101 example (date)*

View the end-user version of the report



Google Cloud

View your report as an end-user



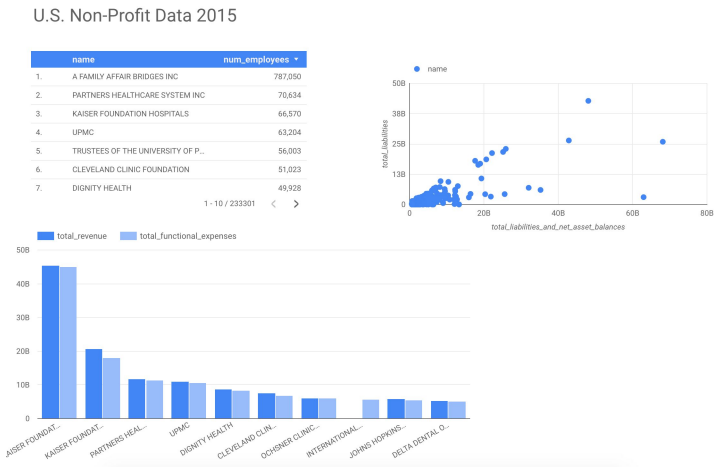
Google Cloud

And here is your report. Notice it looks very similar to when you were editing it, but as a viewer, you can't modify the report.

Try mousing over the chart: notice it is live data, not just a static image.

So this report may not look like much, but with just a few mouse clicks, you connected to a Google Analytics data source and created a chart graphing session data over time.

Demo: Google Data Studio Walkthrough



Ecommerce Dashboard

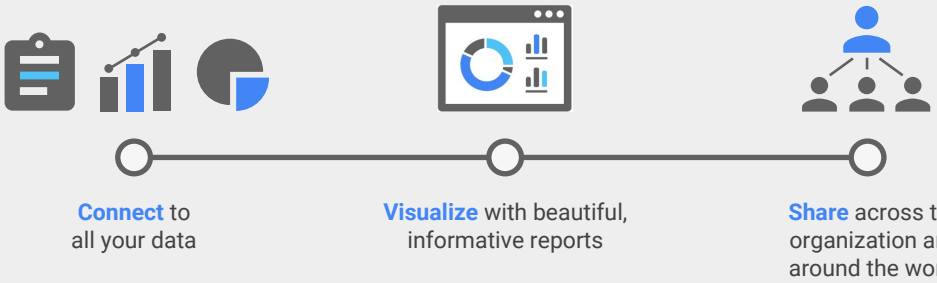
Google Cloud

Access the Data Studio homepage: https://datastudio.google.com/c/u/0/

Ecommerce Dashboard to Walk through
https://datastudio.google.com/c/u/0/org/UTgoe29uR0C3F1FBAYBSww/reporting/0B2-rNcnRS4x5UG50LTBMT0E4aXM/page/nQN

Or build IRS dataset from BigQuery

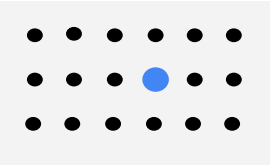
Use Data Studio to explore and share your data insights



Google Cloud

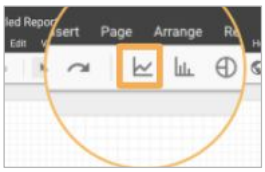
Data Studio's vision as a product is to simplify each of the critical steps in producing reports and dashboards. The Data Studio mantra is "Connect, visualize, share."

Summary: Explore and present your insights visually



Guide the eye of your user with preattentive attributes

Use the right visual to convey the right message



Click-and-drag new report charts inside Data Studio

Google Cloud

Visualizing data is both an art and a science. We've just scratched the surface of visualization theory when we discussed pre-attentive vs post-attentive processing for quick eye-to-brain understanding. Along the way we saw some terrifically bad ways to visualize data with the wrong charts and picked up some best practices along the way.

Lastly, we looked at Google Data Studio which is the visualization platform you will be exploring in more depth in your next lab.

Images from Data Studio public demo:
https://datastudio.google.com/c/u/0/org/UTgoe29uR0C3F1FBAYBSww/reporting/0B2-rNcnRS4x5UG50LTBMT0E4aXM/page/nQN

Lab 7

Exploring a Dataset in Google Data Studio

Google Cloud

Exploring a Dataset in Google Data Studio

In this lab, you will visually explore Google BigQuery data tables inside of Google Data Studio. You will look for relationships and insights between fields in your dataset.