

+tableau®

INTRO TO TABLEAU

TA: Tanner Finken

AGENDA

- What is Tableau
- How does it work?
- 2017 Tennis Data

Example

- Display data, calculated fields, aggregates and joins
- Exporting Results / Concluding Remarks

WHAT IS TABLEAU?

- Data Visualization Software
- Used for Data Analysis and many forms of Data Visualization
- Drag and Drop Interface for faster navigation

SIMILAR PRODUCTS

Power BI (Microsoft)

Others

All use software tools for simple and effective data cleaning and visualization

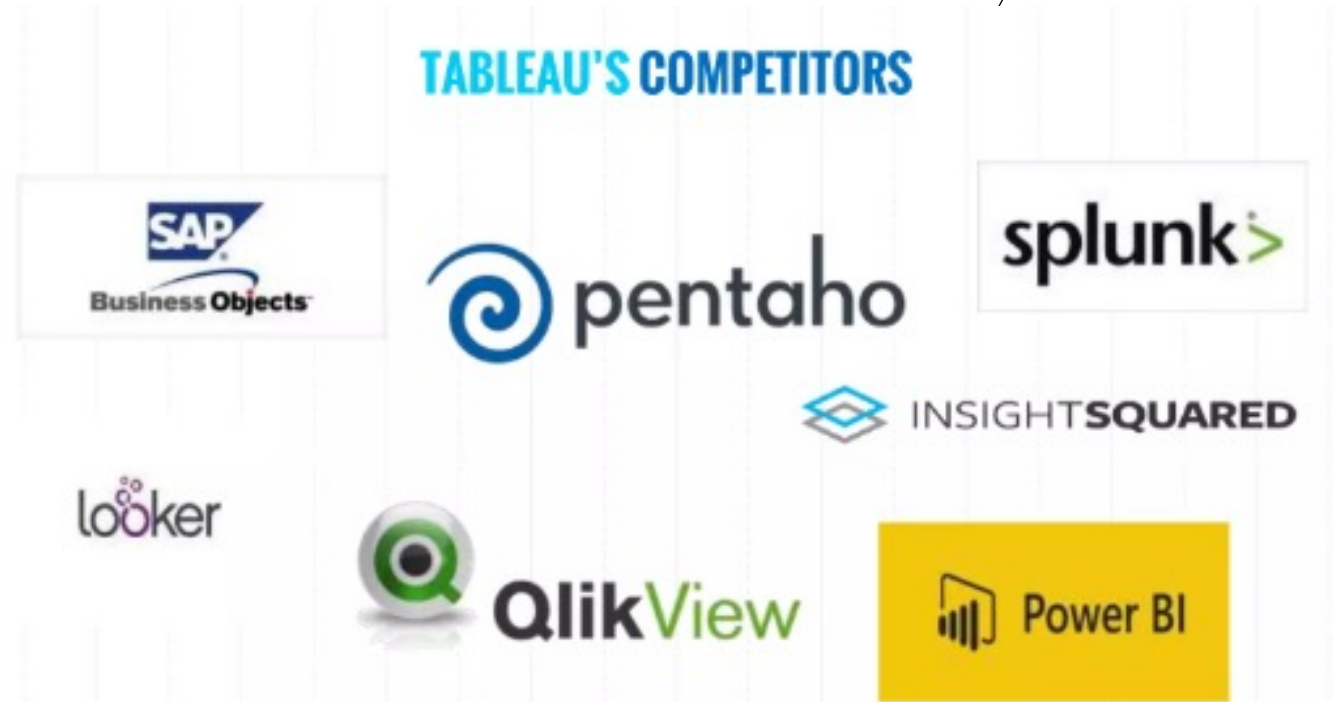
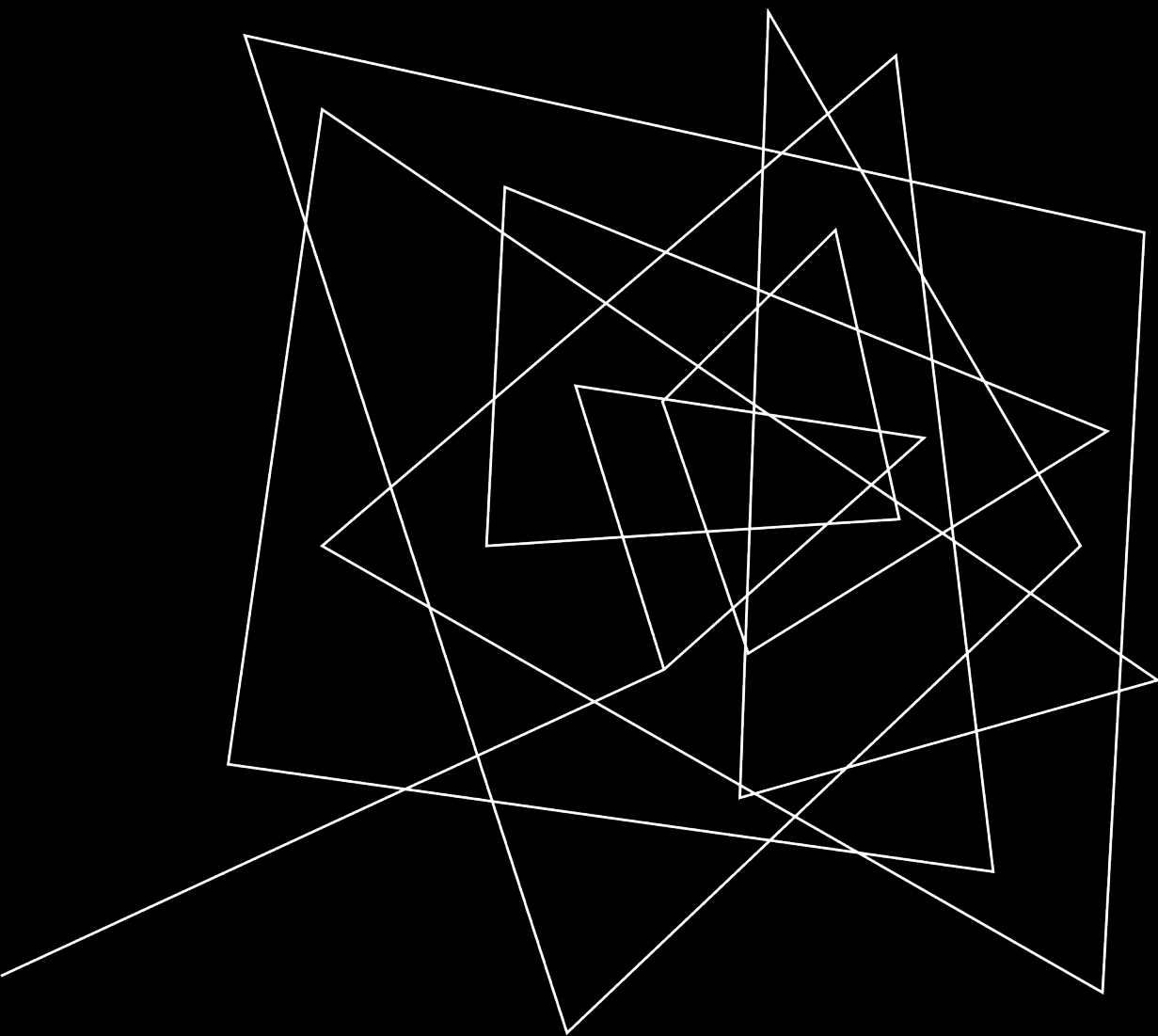


Image from [Tableau Presentation \(slideshare.net\)](https://www.slideshare.net)



HOW DOES IT
WORK?

TYPES OF TABLEAU

Tableau Prep - Data Cleaning Software

Tableau Desktop* – A downloadable application for Data Visualization
(Free for Students)

Tableau Online – Analytics platform fully hosted in the cloud with the ability to publish dashboards and let anyone view them

Tableau Server – an online platform that allows you to host and manage Tableau data sources, workbooks, reports, and dashboards

Tableau Reader – Free software tool that allows you to view workbooks and visualizations made in Tableau Desktop or Public

Tableau Public – Free Visualization Software with Limitations (Visuals can only be published and not saved or exported locally)

***- Tableau Desktop is the one we will be learning about**

TABLEAU DESKTOP DOWNLOAD

Tableau Desktop ([free for students](#) for 1 year)

Fill out [this form](#)

Download Tableau Desktop [here](#)

Use the email code to verify, then register and begin using

You're almost there!

Are you an instructor? Visit tableau.com/teaching to request your license.

Students at accredited academic institutions worldwide are eligible for a free one-year license to activate Tableau Desktop and Tableau Prep. Complete the form below to confirm your eligibility and unlock your new free license. You must be 16 years of age or older to request a license. [How does verifying work?](#)

Country/Region

Country/Region (of school)*

Select One

Personal information

Legal First Name*

Legal First Name

Legal Last Name*

Legal Last Name

Email*

Email

Confirm Email*

Confirm Email

Date of birth*

Month Day Year

salesforce

tableau

Why Tableau Products Solutions Resources Partners

PRODUCT RELEASES

Tableau Desktop 2022.3

We recommend using the newest maintenance release of this version, which contains additional fixes.

DOWNLOAD TABLEAU DESKTOP 2022.3

DOWNLOAD LATEST VERSION (RECOMMENDED)

← BACK TO ALL VERSIONS

Intro To Tableau

Activate Tableau

Activation

Activate the product.

Enter product key:

0000-0000-0000-0000-0000

If you need help with your product key, please file a support request at <http://www.tableau.com/support/request>.

< Back Activate

DATA

Import your data from a variety of data sources (SQL, Oracle, Salesforce, JSON, CSV file, etc)

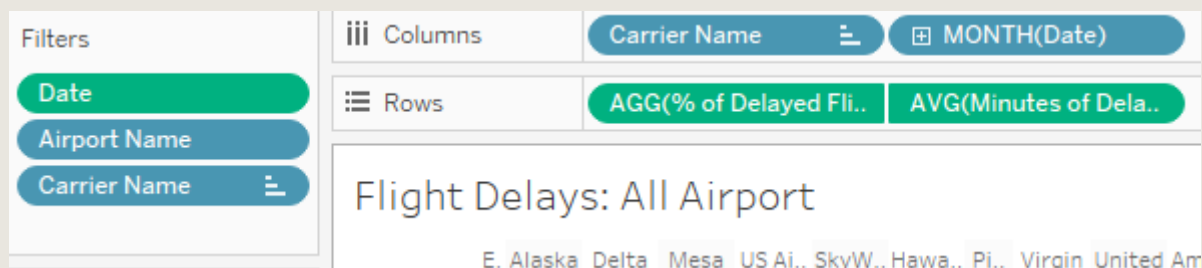
Data is assessed for type automatically and given default assumptions

Blue = Categorical or String Data types (Called Dimensions)

Green = Numeric Data (Called Measures)

These are just the defaults and can be changed as you are crafting your visual

Each visual tells you what it needs to be crafted



For **side-by-side circles** try

1 or more **Dimensions**

1 or more **Measures**

Requires at least 3 fields

FORMAT

Here is what the Tableau editor consists of
It is a lot, but each part has its own task

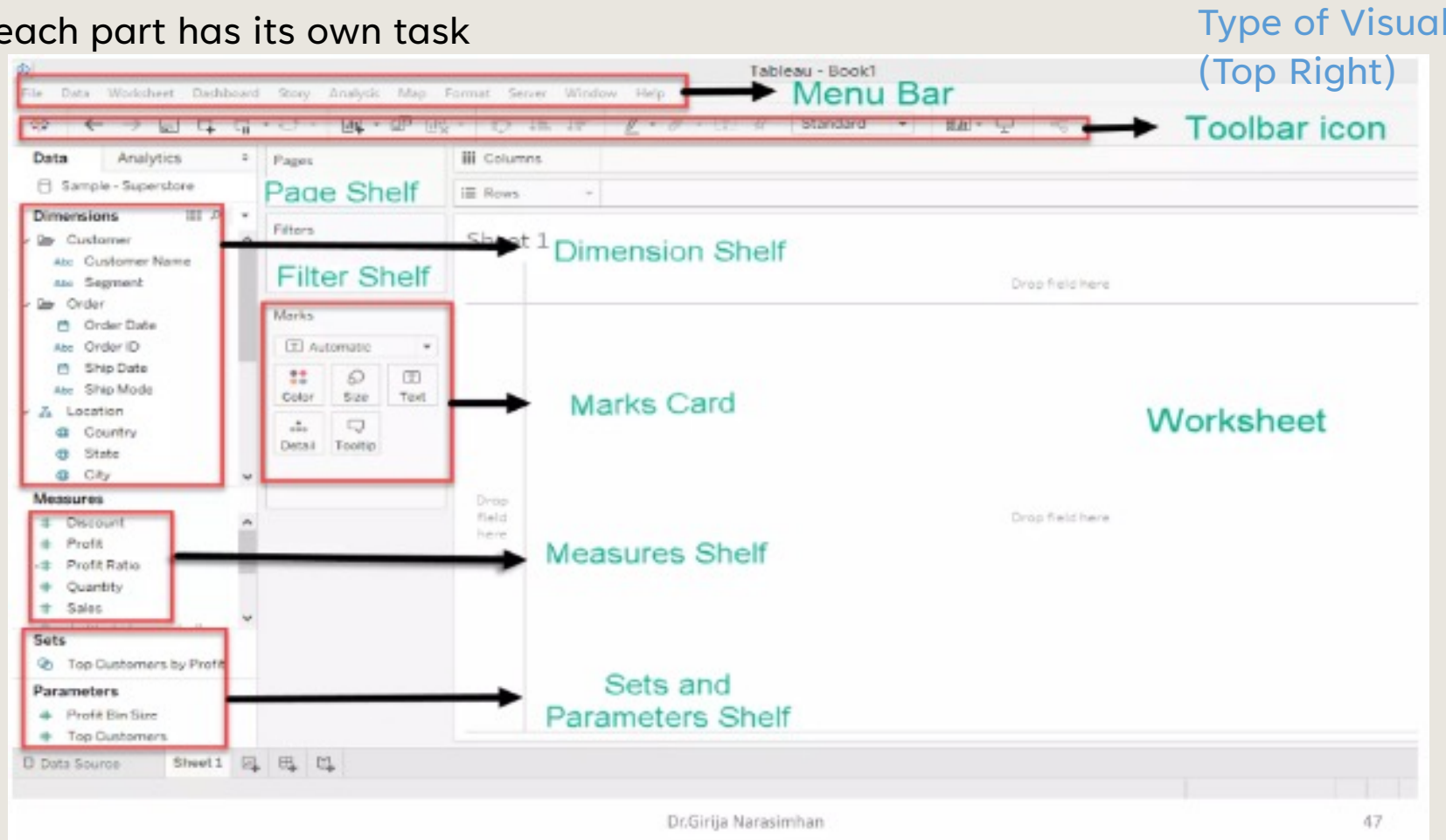
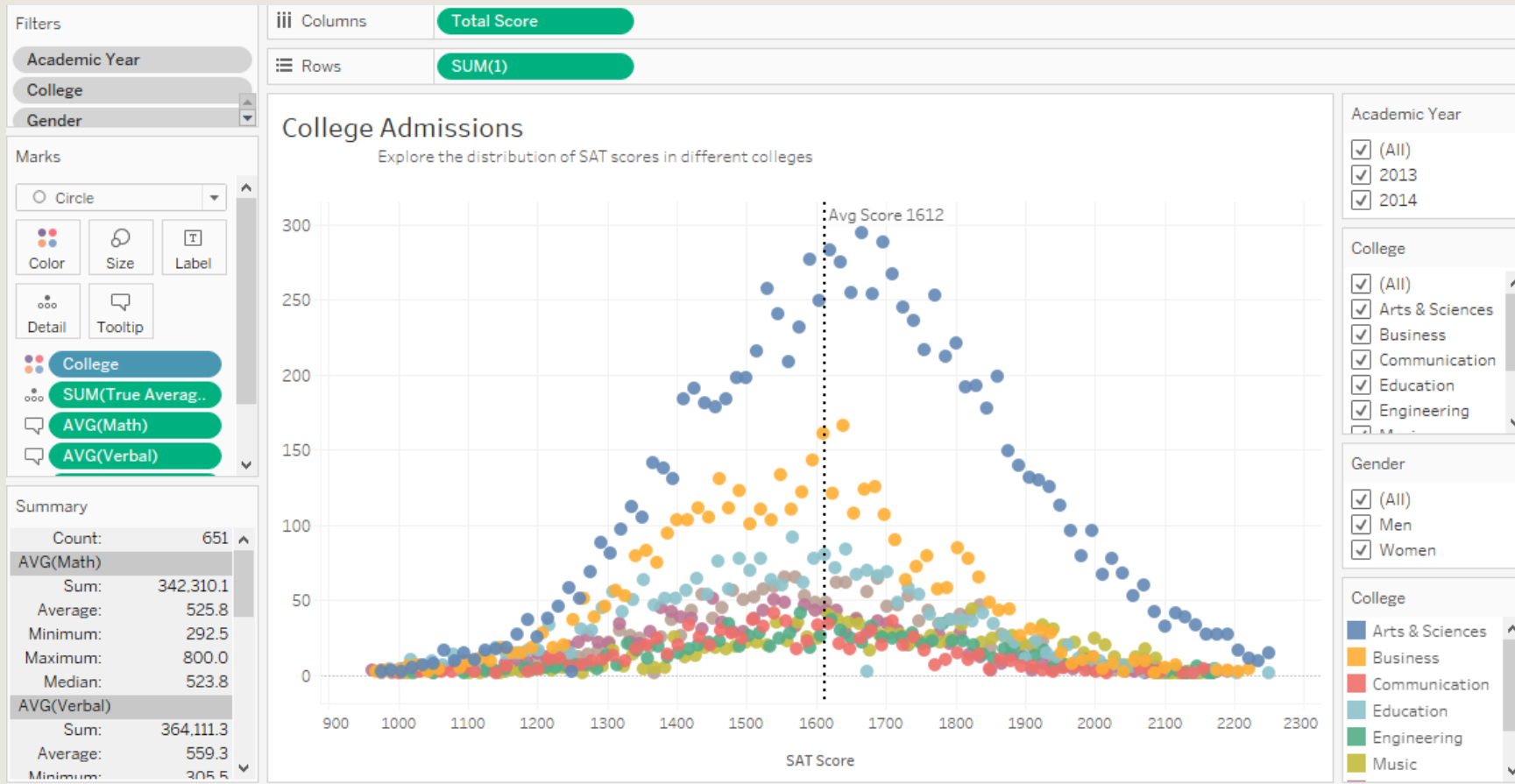
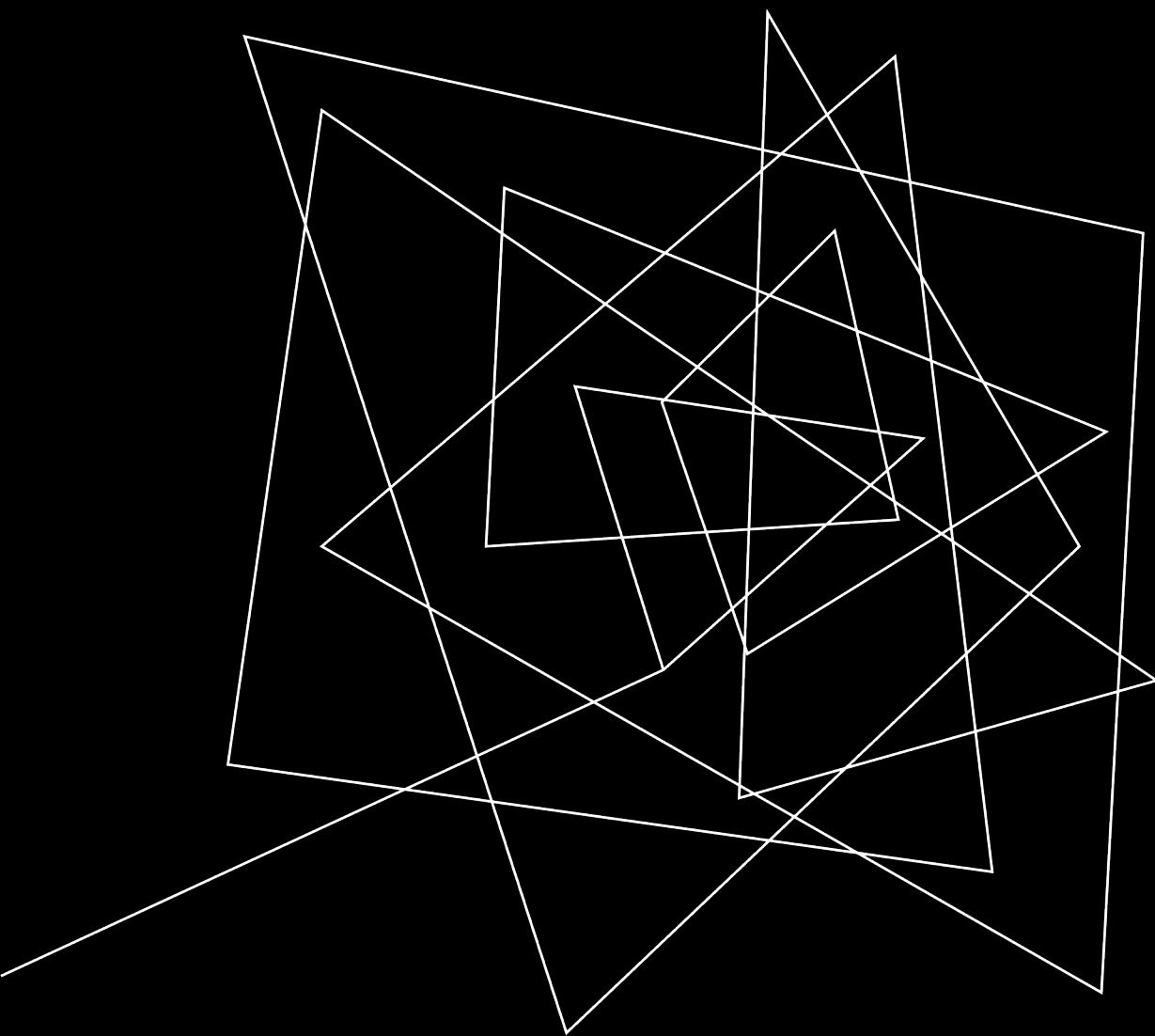


Image from [Visualization using Tableau \(slideshare.net\)](https://www.slideshare.net/Visualization-using-Tableau)

GOAL

Visualize data quickly and be able to make multiple graphs into a dashboard to answer one or multiple questions





**EXAMPLE:
2017 MEN'S
TENNIS DATA**

DATA

The data was found at the following website: [ATP World Tour tennis data - Dataset - DataHub - Frictionless Data](#)

In particular, we will be using the 2017 match_scores and match_stats data found here:

[https://datahub.io/sports-data/atp-world-tour-tennis-data/r/match_scores_2017_unindexed.csv](#)

And here:

[https://datahub.io/sports-data/atp-world-tour-tennis-data/r/match_stats_2017_unindexed.csv](#)

The easiest way is to just download them to your computer

STARTING THE PROJECT

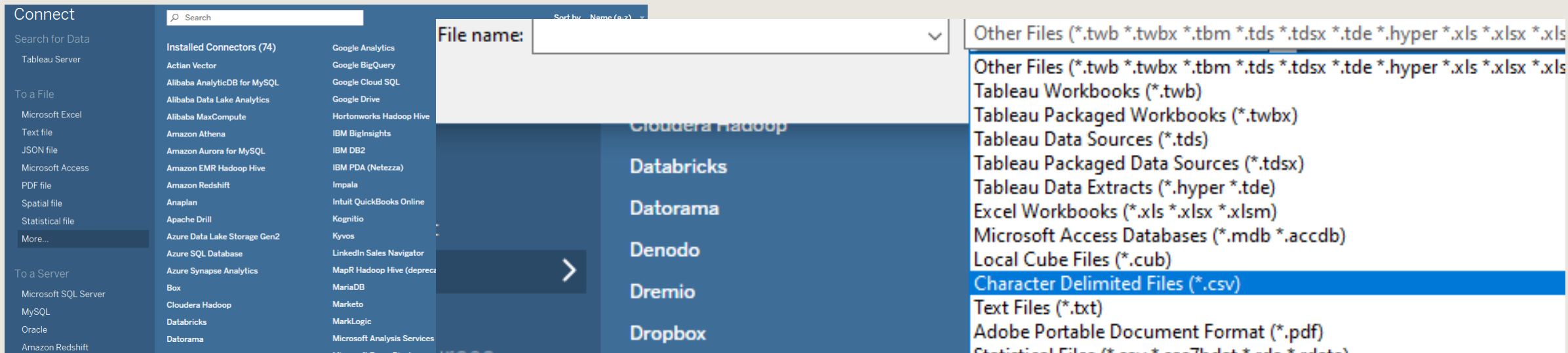
Create a new workspace: File > New (Ctrl-N)

Connect to the data source: Data > New Data Source (Ctrl-D)

Under File Source, select more, then CSV file type and select one of the tennis match files.

Then under data source, add another connection and add the other CSV file

Finally, connect them together using Match Id



VIEWING DATA ATTRIBUTES

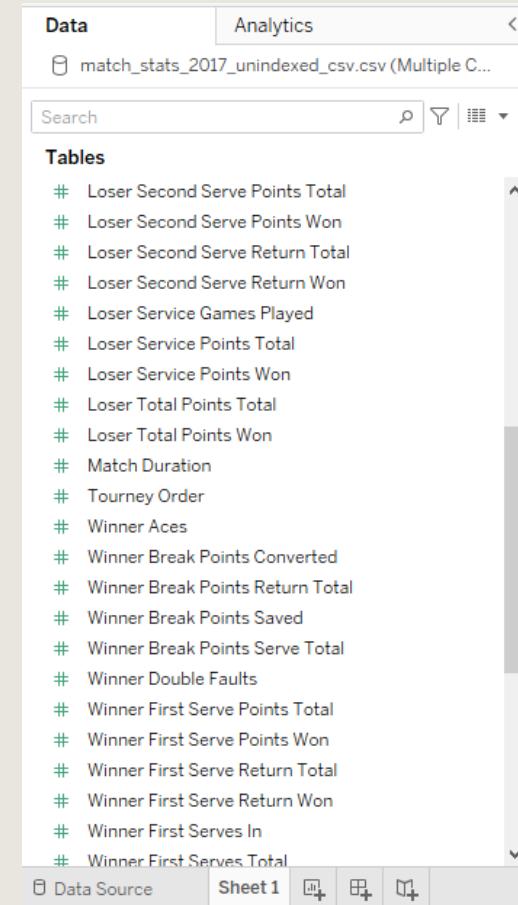
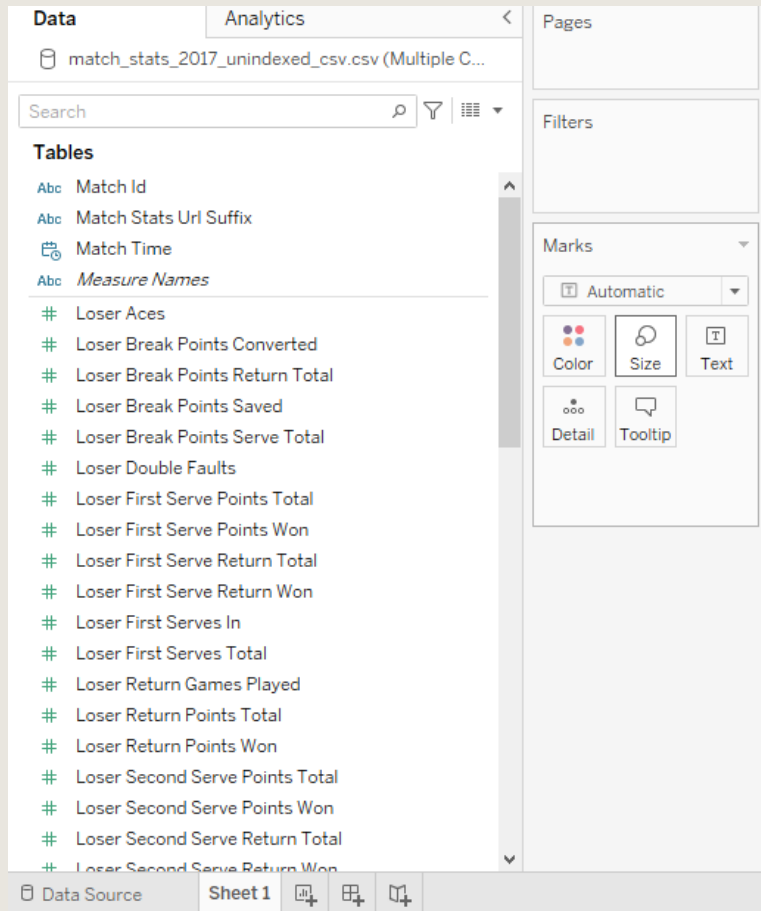
First, we need to take a look at the data and see what we can visualize with the information we have. Once the data is connected, Under the Data Source tab at the bottom, we can see what attributes the data has

# match_scores_2017_unindexed_csv.... Match Order	Abc match_scores_2017_unindexed_csv.... Winner Name	Abc match_scores_2017_unindexed_csv.... Winner Player Id	Abc match_scores_2017_unindexed_csv.... Winner Slug	Abc match_scores_2017_unindexed_csv.... Loser Name
1	Grigor Dimitrov	d875	grigor-dimitrov	Kei Nishikori
1	Grigor Dimitrov	d875	grigor-dimitrov	Milos Raonic
2	Kei Nishikori	n552	kei-nishikori	Stan Wawrinka

Abc match_scores_2017_unindexed_csv.... Winner Seed	Abc match_scores_2017_unindexed_csv.... Loser Seed	Abc match_scores_2017_unindexed_csv.... Match Score Tiebreaks	# match_scores_2017_unindexed_csv.... Winner Sets Won	# match_scores_2017_unindexed_csv.... Loser Sets Won	# match_scores_2017_unindexed_csv.... Winner Games Won
7	3	62 26 63	2	1	
7	1	76(7) 62	2	0	
3	2	76(3) 63	2	0	

VIEWING DATA ATTRIBUTES

The other way is to look under the data pane, while editing a sheet



ANSWERING QUESTIONS:

Some questions we will try to answer:

- Who won or lost the most matches in 2017?
- What tournaments had the biggest number of upsets?
- Does the return of serve affect the score of how badly someone loses?

WHO WON OR LOST THE MOST MATCHES IN 2017?:

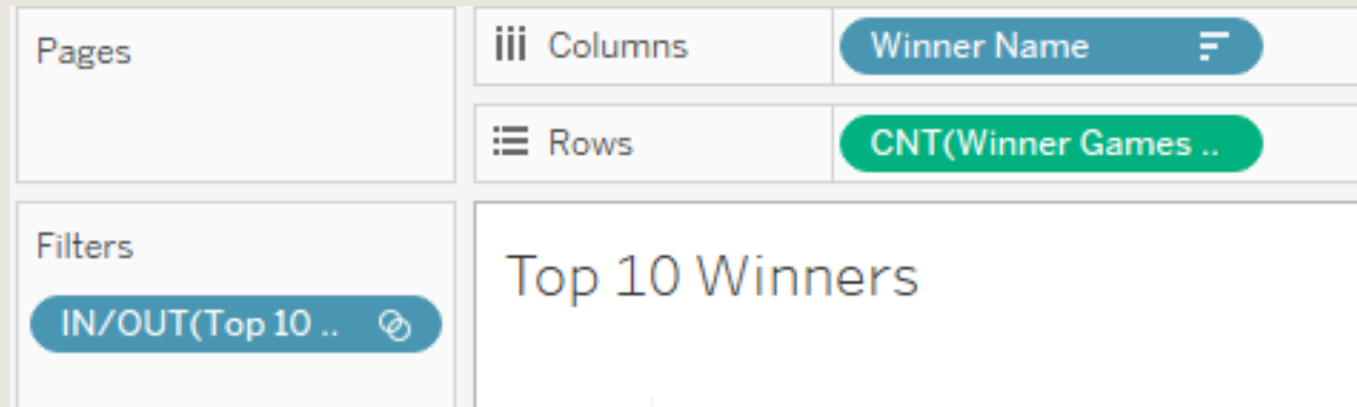
Top 10 Players who won the most matches in 2017

Make a set of the top 10 Winners by right clicking on Winner Name, Create > Set

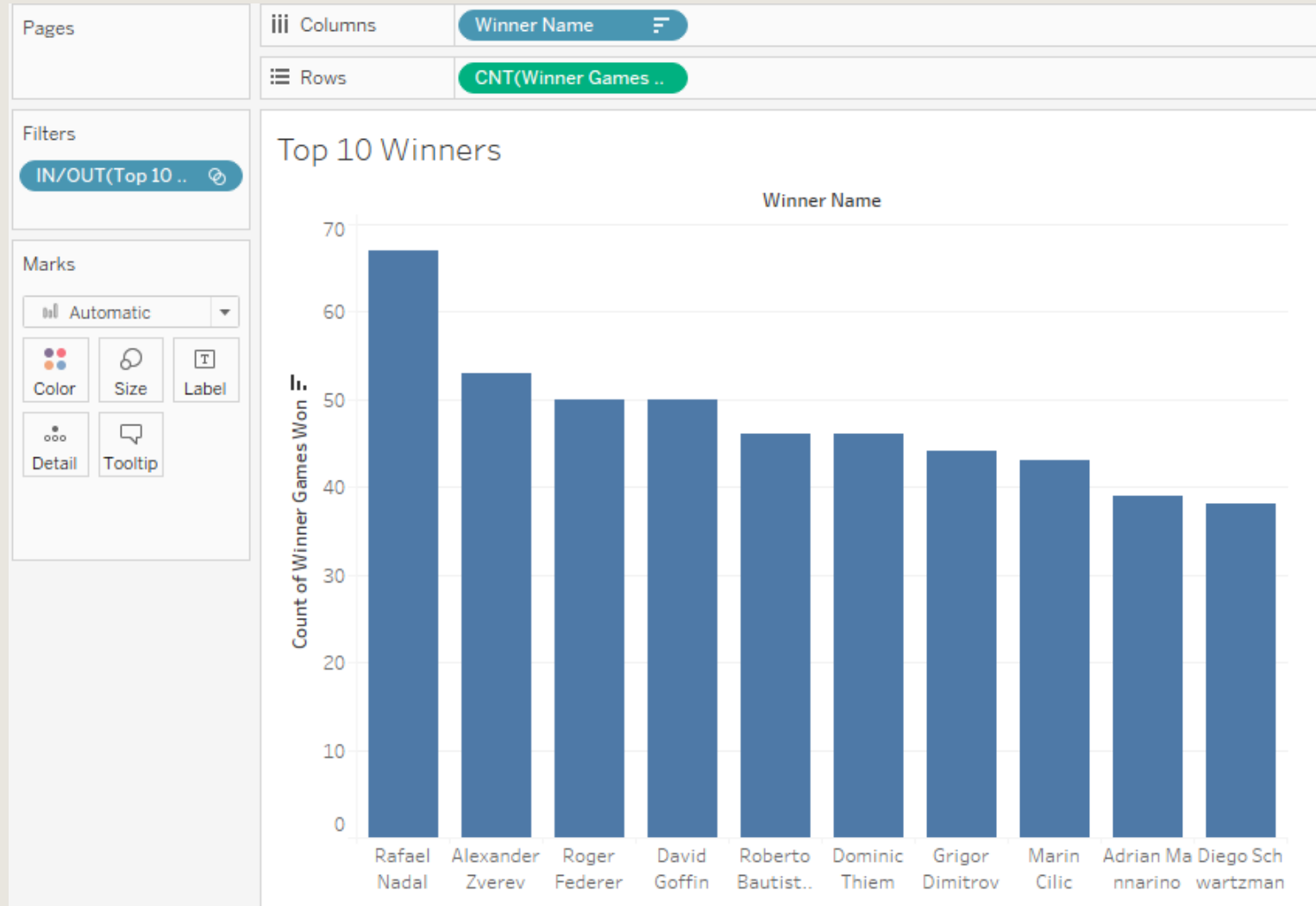
Change name to “Top 10 Winners”, then under the Top tab, we change it to top 10 by Winner Games (Count)

Simply drag Winner Name into the Column and then drag Count (change Aggregation type) of Winner Games Won into the Row

Then filter using the set we created



WHO WON OR LOST THE MOST MATCHES IN 2017?:



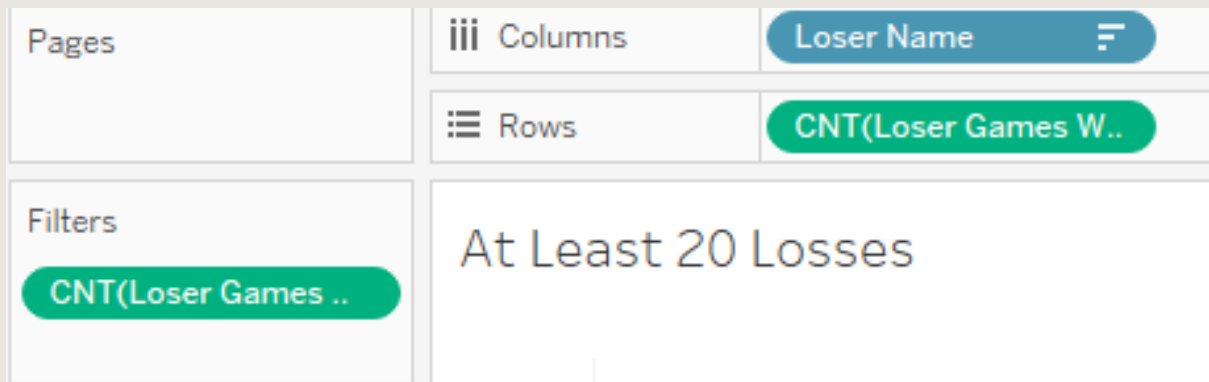
WHO WON OR LOST THE MOST MATCHES IN 2017?:

Players who lost more than 20 matches in 2017

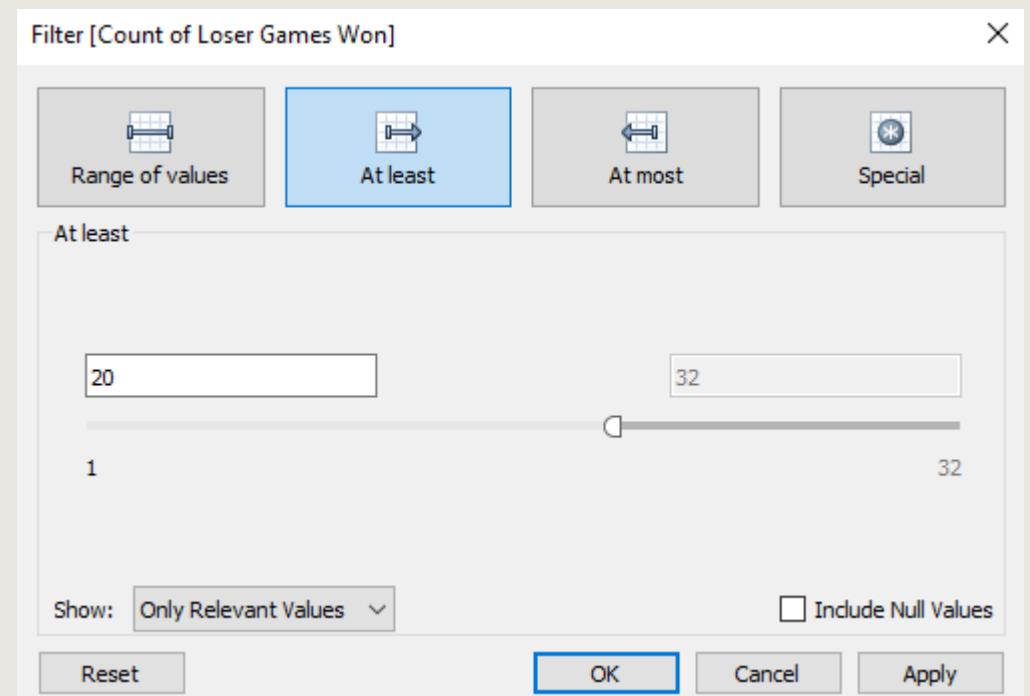
Create new sheet 

Simply drag Loser Name into the Column and then drag Count (change Aggregation type) of Loser Games Won into the Row

Then filter based on the number of Loser Games Won, at least 20

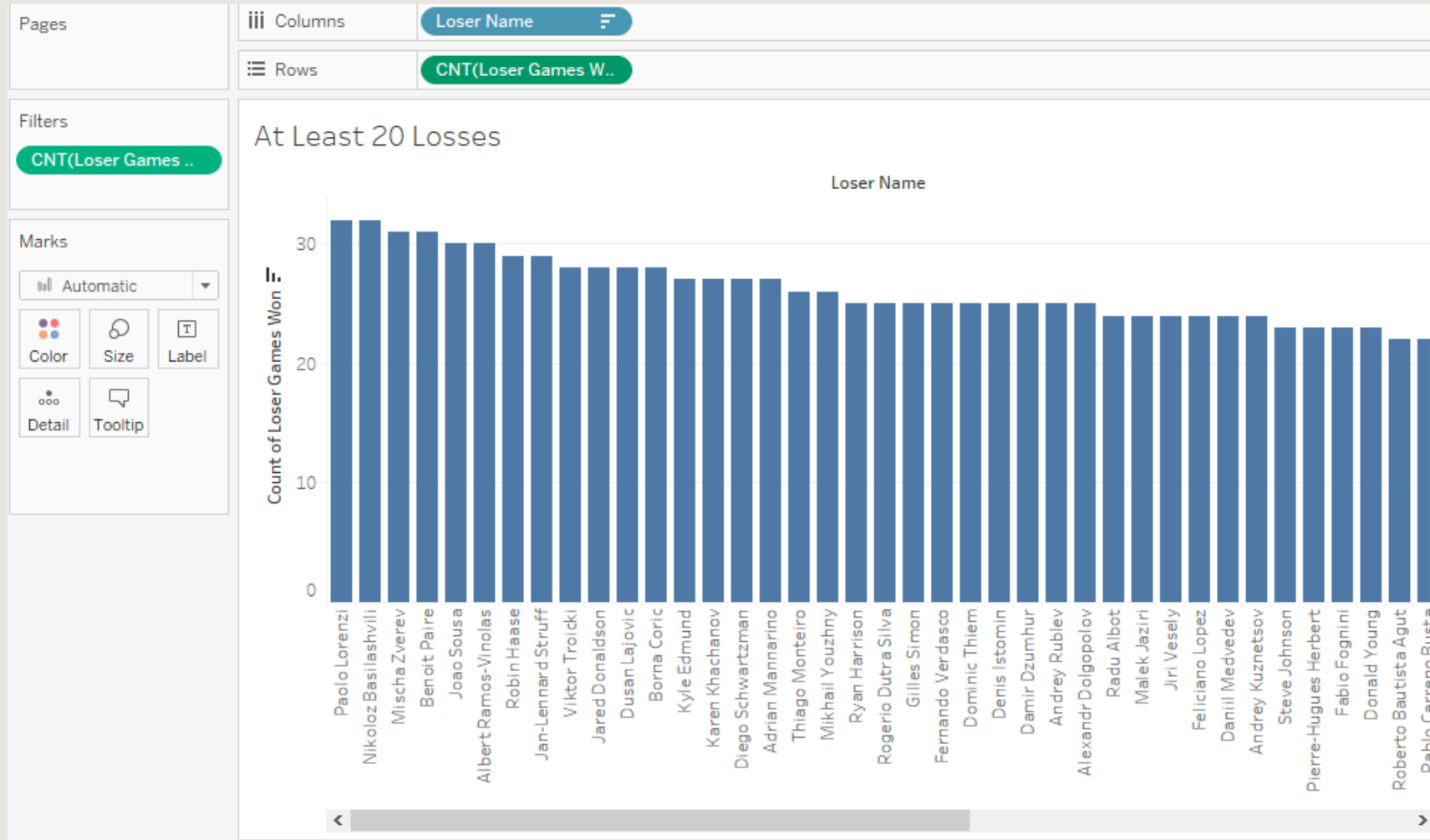


The screenshot shows the Tableau interface. On the left, the 'Columns' shelf contains 'Loser Name' and the 'Rows' shelf contains 'CNT(Loser Games W..)'. Below the shelves, a filter card is visible with the text 'At Least 20 Losses'. The main visualization area is currently blank.



The screenshot shows the 'Filter [Count of Loser Games Won]' dialog box. The 'At least' tab is selected. The 'At least' section shows a range from 20 to 32. The 'Show:' dropdown is set to 'Only Relevant Values' and the 'Include Null Values' checkbox is unchecked. The 'OK' button is highlighted.

WHO WON OR LOST THE MOST MATCHES IN 2017?:



WHAT TOURNAMENTS HAD THE BIGGEST NUMBER OF UPSETS?

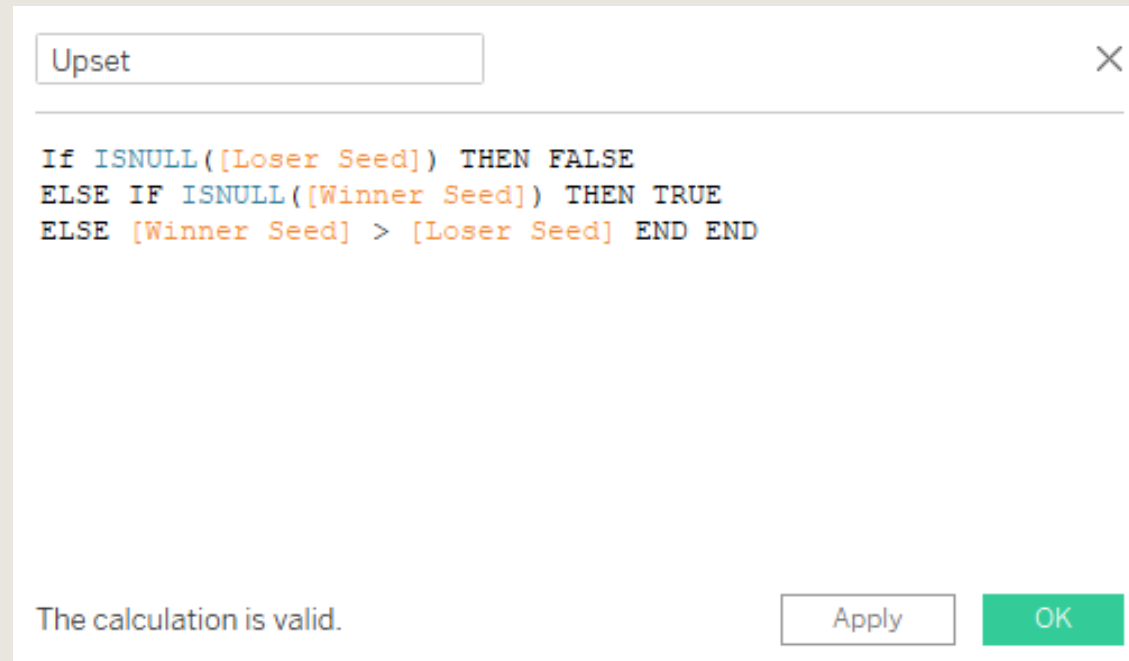
Define an upset: when a lower seeded player beats a higher seed (Lower value of seed is better e.g. 1 is the best seed)

First, create a calculated field, seeing where upsets occur (Winner Seed > Loser Seed)

Then, we have to handle null values in the data, usually null means the player is unranked.

If a loser is unranked, it's not an upset. If the winner is unranked, it is

So, we get the following formula:



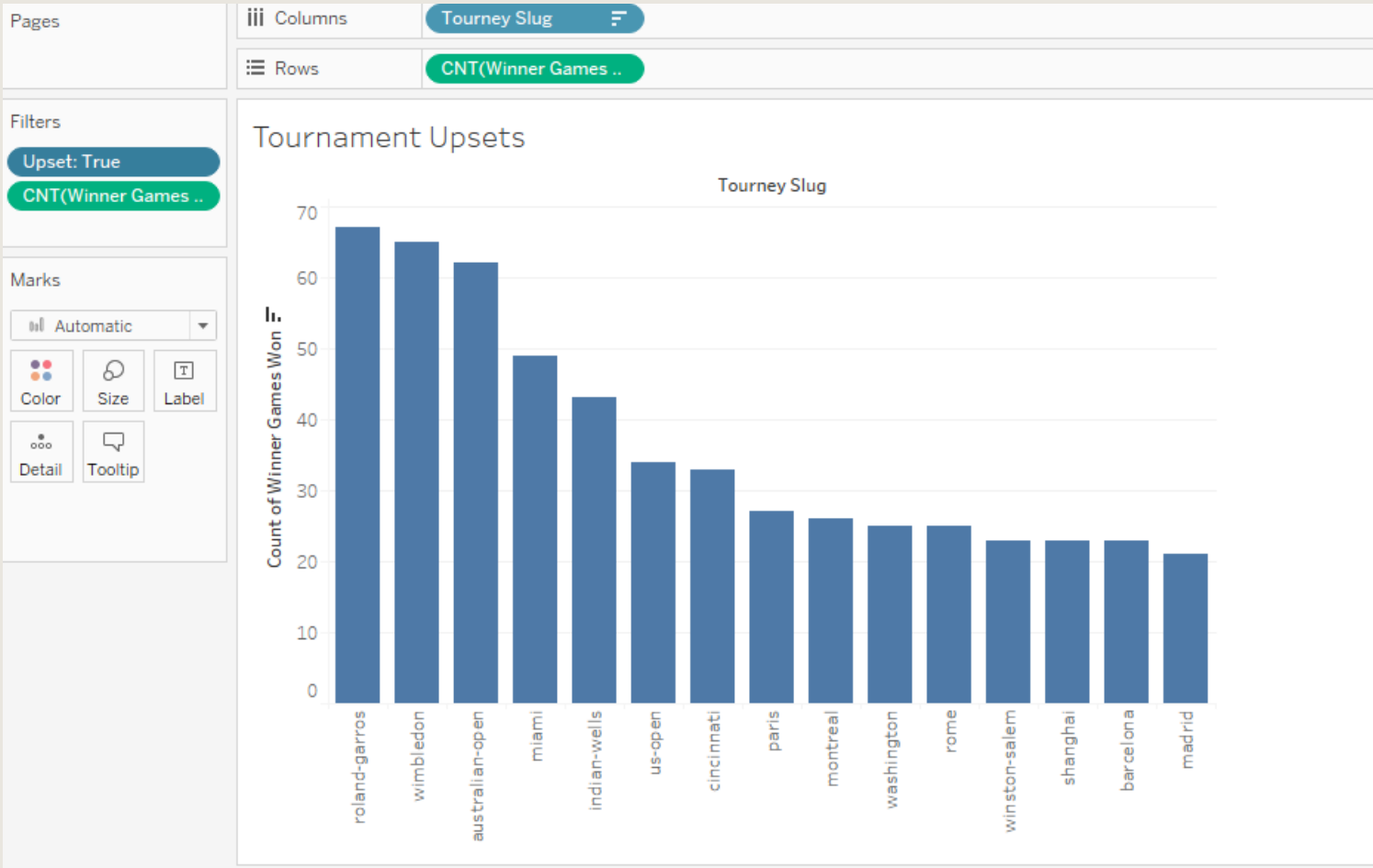
WHAT TOURNAMENTS HAD THE BIGGEST NUMBER OF UPSETS?

Then, drag Tournament names to columns, and (Count) of Winner Games Won(Doesn't matter what field)
Filter based on the newly created upset field, and by the count if you don't want all tournaments

The screenshot shows a Tableau interface. On the left, there are two filter cards: 'Upset: True' and 'CNT(Winner Games ..'. The main view displays 'Tournament Upsets' with a value of '70'. The columns shelf contains 'Tourney Slug' and the rows shelf contains 'CNT(Winner Games ..'.

The screenshot shows the 'Filter [Upset]' dialog box. It has three tabs: 'General', 'Condition', and 'Top'. The 'General' tab is selected. There are three radio buttons: 'Select from list' (selected), 'Custom value list', and 'Use all'. Below this is a search box with the text 'Enter search text'. There are two checkboxes: 'False' (unchecked) and 'True' (checked). At the bottom, there are buttons for 'All', 'None', and 'Exclude' (unchecked). A 'Summary' section shows: Field: [Upset], Selection: Selected 1 of 2 values, Wildcard: All, Condition: None, Limit: None. At the bottom of the dialog are buttons for 'Reset', 'OK', 'Cancel', and 'Apply'.

WHAT TOURNAMENTS HAD THE BIGGEST NUMBER OF UPSETS?



DOES THE RETURN OF SERVE AFFECT THE SCORE OF HOW BADLY SOMEONE LOSES?

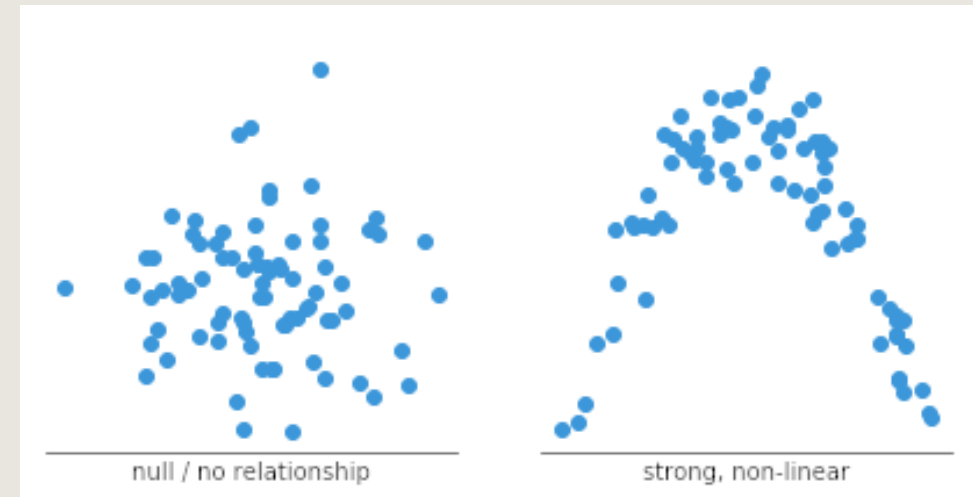
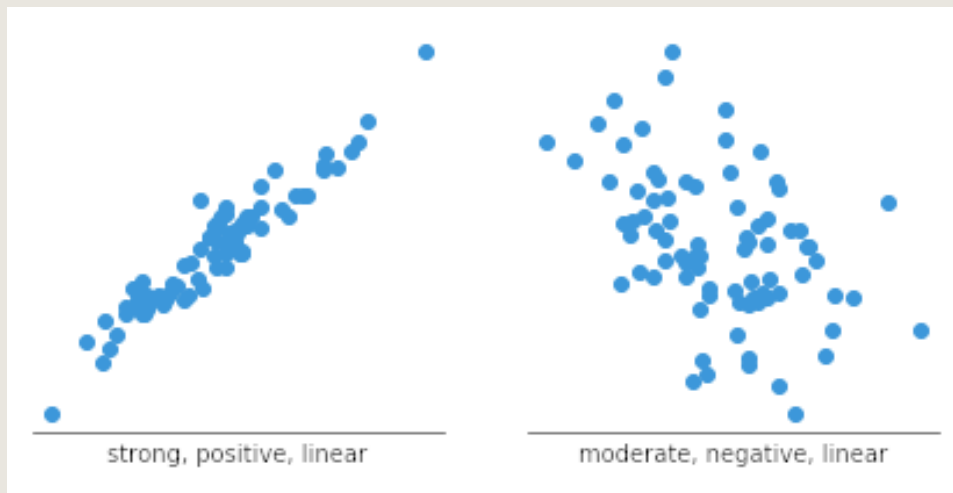
Define how what we measure for return of serve:

In the data, we have loser return points won, loser first serve return won, and loser second serve return won. (All are values that should be converted to percentages by dividing by total)

Create those calculated fields: $[\text{Return won}] / [\text{Return total}]$

Then, measures of how badly someone loss can be how many games they were able to win, how long the match was played, or how many points they won.

So, we can visualize these using scatterplots



DOES THE RETURN OF SERVE AFFECT THE SCORE OF HOW BADLY SOMEONE LOSES?

Drag all the percentages we just calculated into columns
Then, drag the values we want to compare against into rows
And filter out any null values in the data

The screenshot shows the Tableau interface with the following components:

- Columns shelf:** Contains three green pills: "Loser Return Points ..", "Loser First Serve Ret..", and "Loser Second Serve ..".
- Rows shelf:** Contains three green pills: "Loser Games Won", "Match Duration", and "Loser Total Points W..".
- Filters shelf:** Contains one green pill: "Loser Second Serve ..".
- View:** A visualization titled "Returning Stats" is partially visible at the bottom.

Special Values for [Loser Second Serve Return Percent] [X]

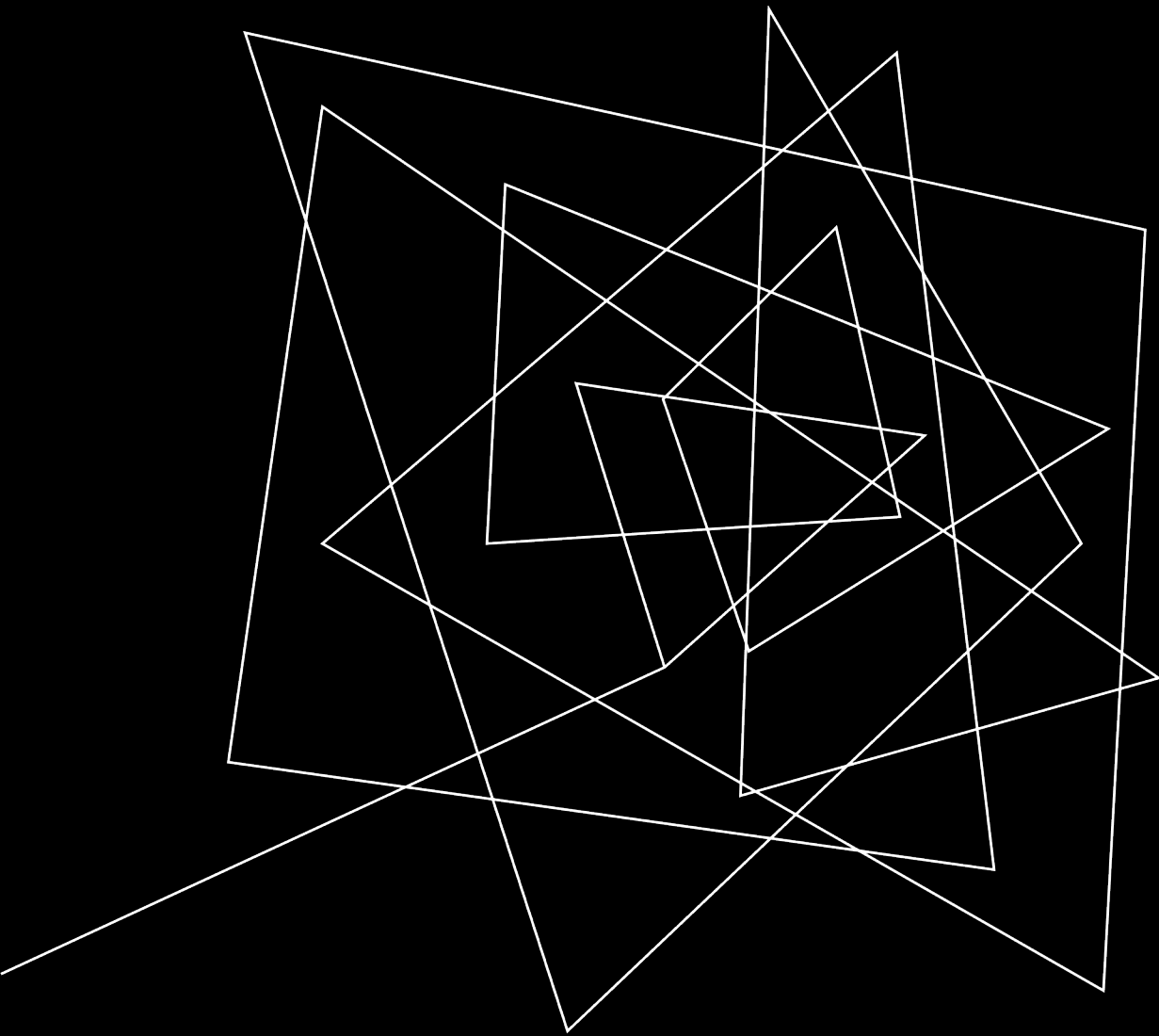
The **Loser Second Serve Return Percent** axis has at least 11 null values.
What do you want to do?

Filter data
Exclude the special values from the view and calculations.

Show data at default position
Show the special values at a default position on the axis. For example, Null values are shown at 0.

DOES THE RETURN OF SERVE AFFECT THE SCORE OF HOW BADLY SOMEONE LOSES?





EXPORTING
RESULTS/
CONCLUDING
REMARKS

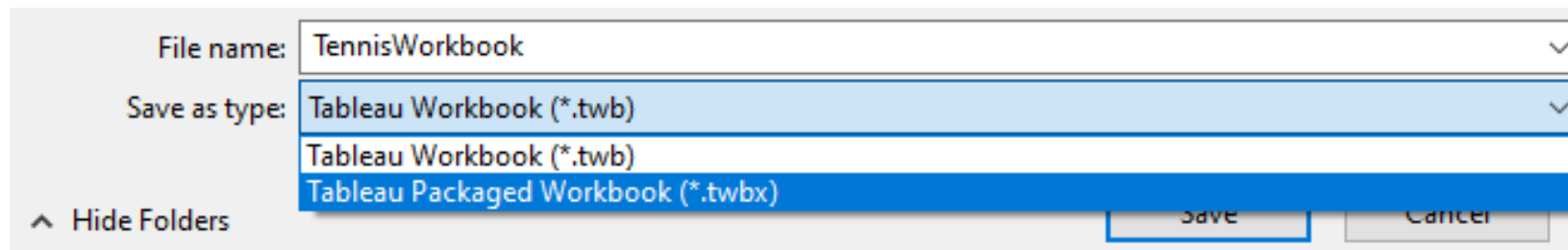
SAVING YOUR WORKBOOK

You can save your graphs for later editing as Tableau Workbook (.twb) or Tableau Packaged Workbook (.twbx)

The main difference is that the Packaged Workbook includes your data along with the graphs, so you can run it on another computer without needing access to the data.

While a Workbook is quicker to save and edit

Simply use File > Save As



EXPORT TO IMAGE OR OTHER FORMATS

You can export your graphs to an image by using

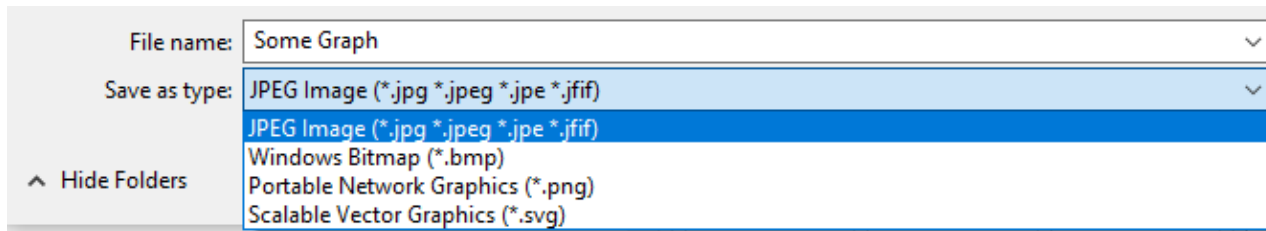
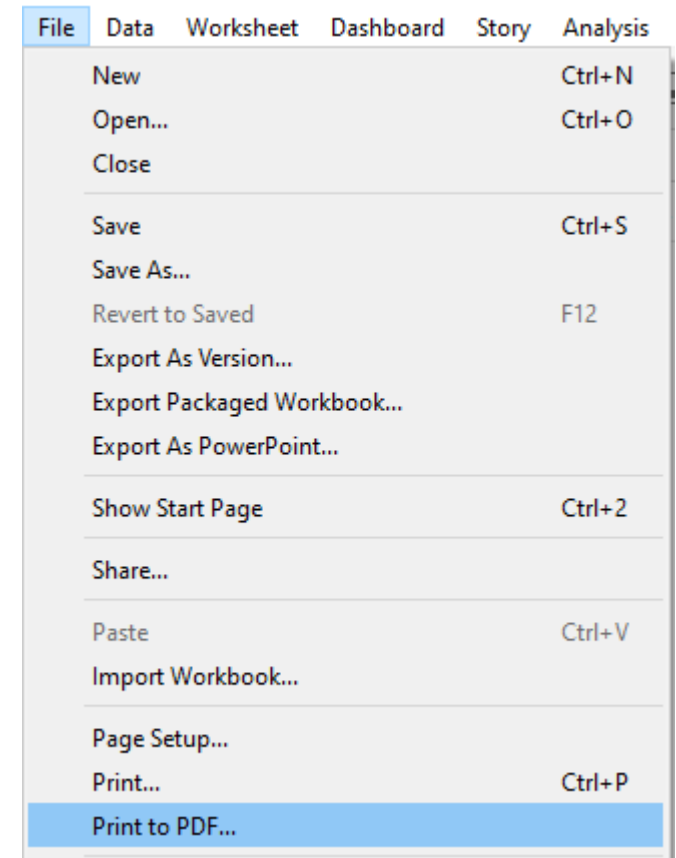
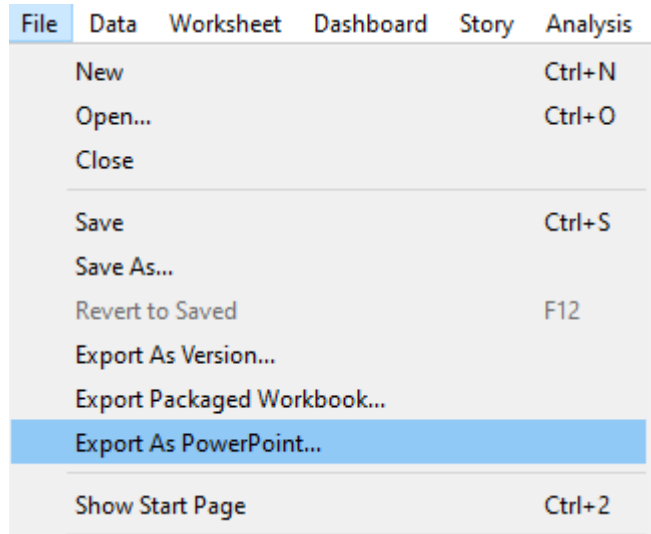
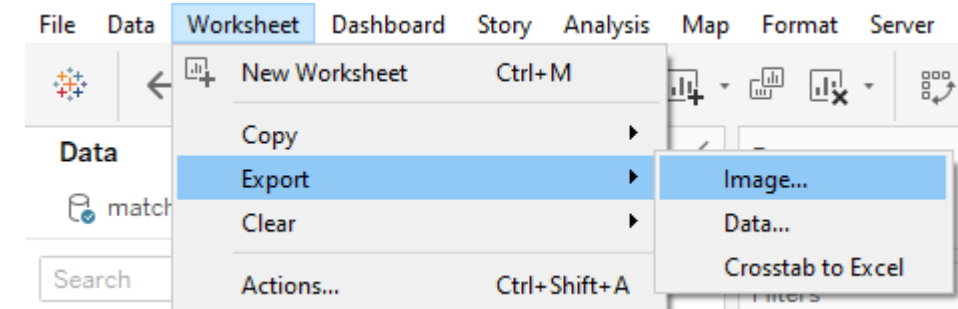
Worksheet > Export > Image

Or you can download as a PowerPoint with

File > Export as PowerPoint

Or as a PDF

File > Print to PDF



Two thin black lines intersect on the left side of the slide. One line is horizontal, and the other is diagonal, crossing it from the top-left towards the bottom-right.

CONCLUDING REMARKS

Tableau is a quick and effective data visualization tool, which can be used to quickly analyze and see data. The main drawbacks are taking time to learn how to use the tools that Tableau provides and getting used to the drag and drop style of building graphs. Creating a basic graph usually takes a lot less time than coding libraries such as D3 and Vega but can require more time with online tutorials to change small parts. Although not shown here, Tableau offers a wide variety of graphs that can be generated and is very useful for visualizing many types of data. Then, the visuals can be put into a dashboard for easy access in the same spot.



THANK YOU

Tanner Finken

finkent@arizona.edu