# **Tableau Handout Solutions**

## **Table of Contents**

1
2
2
3
3
7
14

# **Tableau Basics**

**Tableau** is used to generate interactive data visualizations and makes very common visualizations very easy. Our focus today is on creating simple visualizations using Tableau.

In this tutorial you will learn:

Data	Know how to import data
	View and describe data within Tableau
	Sort Data
	Select data types (Date, Number, String)
	Make Calculated Fields
	Change the # of rows displayed
	Create Groups
Sheet	Differentiate measure and dimensions
	Create simple visualization and change chart type
	Adjust encodings
	Adjust color
	Adjust aggregation methods (e.g. sum, average, count)
	Sort axis
Organize Sheets	Build dashboard
	Build stories
Action	Create annotations on the chart
	Create data filters
	Link sheets together with actions.
	Add analytics enhancements (e.g., constant line, regression)
	Edit legend and filter by legend.

# **Tableau Setup & Resources**

Please make sure you have Tableau installed and running. You can get a copy at: <a href="https://www.tableau.com/academic/students">https://www.tableau.com/academic/students</a>

Make sure you can run Tableau. If you need more information on Tableau take a look at:

- Videos:
  - o http://www.tableau.com/learn/training
- Books:
  - o <u>https://link.springer.com/book/10.1007/978-1-4842-1934-8</u>
- Other Recommended Resources:
  - o Youtube!

# **Tableau Terms**

Sheet: A sheet is a singular chart/map in Tableau.

**Dashboard:** A dashboard is a canvas for displaying multiple sheets at a time and allowing them to interact with each other.

Workbook: A workbook is the entire Tableau file containing your sheets and dashboards.

**Measure:** A variable from the dataset that is meant to be aggregated. (This means it should be a number that it makes sense to do math with: sum, average, and so on.) Measures are often continuous data. Examples include GPA, sales, quantity, quota, height, and salary.

**Dimension:** A categorical variable from the dataset that is used to slice and dice the data into different categories. Dimensions are often discrete data. Examples include country, gender, student ID, and name.

**Rows and Columns Shelves:** The Rows shelf and the Columns shelf is where you determine which variables will go on what axis. Put data you want displayed along the X-axis on the Columns shelf and data you want displayed on the Y-axis on the Rows shelf.

**Filter:** A filter is used to limit what data is being displayed on the sheet. Visible controls for a filter on a sheet or dashboard are called Quick Filters.

**Tooltip:** Tooltips are text boxes that appear when hovering over a mark on a sheet in order to give more information. The text and text formatting in them are easily edited through the Marks card.

**Marks card:** The Marks card is the tool used to create a sheet that controls most of the visual elements in a sheet. Using the Marks card, you can switch between different chart types (bar, line, symbol, filled map, and so), change colors and sizes, add labels, change the level of detail, and edit the tool tips.

# **The Movie Dataset**

**Scenario**: you want to invest in the movie industry. However, you are not sure which director, which movie, or what movie genre that you want to invest your money in. In order to make a wise investment, you obtained some data about movies. You want to create visualizations to help you to make decisions.

### PART 1: Data set-up and manipulation

Download the movies dataset <u>here</u> with your Stanford credential, open the Tableau app, and load the data.

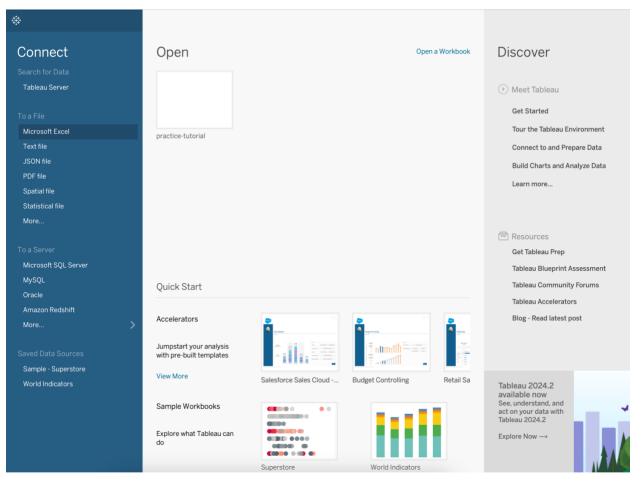


Figure 1

You can load in the data by either clicking to "Open a Workbook" or clicking the "Microsoft Excel option." Once you load it, you will see something like this:

$ \  \  \ast \  \  \leftarrow \  \  \rightarrow \  \  \leftarrow \  \  \  \  \  \  \  \  \  \  \$	0- m	ovies.csv (tableau_	tutorial	Connection <ul> <li>Live</li> </ul>	) Ext	ract	Filters 0 Add
Connections Add							
tableau_tutorial_dataset Microsoft Excel	movies	.csv					
Sheets $ ho$							
Use Data Interpreter							
Data Interpreter might be able to clean your Microsoft Excel workbook.							
I movies.csv							
뭠 New Union							
ि New Table Extension							
	movies.c	sv 💌	18 fields 3201	rows		100 →	rows 🕲 🗸
					<	Abc	#
	Name					movies.csv	movies.csv
	movies.cs	57				Title	US Gross
						The Land Girls	14
	Fields					First Love, Last Rites	1
	Туре	Field Name	Physical Table	Remote		I Married a Strange Pe	
	Abc	Title	movies.csv	Title		Let's Talk About Sex	37
	#	US Gross	movies.csv	US Gross		Slam	1,00
	#	Worldwide Gross	movies.csv	Worldwid		Mississippi Mermaid	2
	#	US DVD Sales	movies.csv	US DVD S.		Following	4
Data Source     Revenue as Dimension	Cornedy vs.	Action Comedy vs Action	(2) Sheet 3	Sum and Averag	ge of Rev		ar ⊡, ₽, Ŭ,

Figure 2

You can view and edit the data under the "Data Source" Tab. Make sure that you are displaying all rows in the data, or your results might look a little different from ours. (Figure 2.1)

Show aliases Sho	ow hidden fields 1,000	0 ⇒ rows	Show aliases	Show hidden fields	.201 🔿 rovis
# movies.csv Running Time (min)	Abc movies.csv Distributor	Abc movies.csv Source	ovies.csv roduction Budget	Abc movies.csv Release Date	Abc movies.csv MPAA Rating

Figure 2. 1

*Question 1: Which movie has the highest production budget? Which movie has the lowest production budget?* 

1.1 Sort the dataset by production budget (Figure 3)

#     ▼       movies.csv       Production Budget
200,000
210,000
250,000
300,000
300,000
340,000
350,000



#### Question 2: How do I sort the dataset by releasing date?

2.1 Convert the "Release Date" from a String to a Date. Click the upper left hand corner to access the list of data types. (Figure 4.1, and 4.2)

2.2 Sort the table by "Release Date".

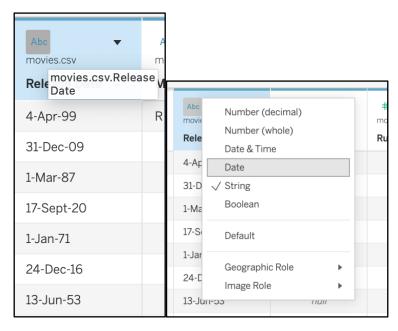


Figure 4.1 and 4.2

**Question 3: Which is the most profitable movie?** 

3.1 Sort the dataset by the "worldwide gross"

3.2 Create a calculated field called "Revenue" that shows the actual profit that these movies make (hint: Revenue= Worldwide Gross-production budget, if we only care about movies that target the international market) (Figure 5)

Evenue	Edit Rename ut Copy Values Hide	Revenue [Worldwide Gross]-[Production Budget]	×
-50,000	Delete ul		
5,128,617 2,900,000 15,050,000	Create Calculated Field		•
7,800,000 4,790,000	Describe Ul	The calculation is valid.	уОК

Figure 5

#### Question 4: How many directors are included in the dataset?

4.1 Click on the triangle next to "Director"-> select "Describe". The "Describe Field" window shows the answer. (Figure 6)

Abc  movies.csv Director	Rename Copy Values	Describe Field	×
Christopher Nolan	Hide	Director	
null	Aliases Create Calculated Field	Role: Discrete Dimension	
null	Create Group	Type: Database column	
null	Split	Remote column: [movies.csv].[Director] Remote type: Unicode character string	
null	Custom Split	Contains NULL: Unknown	
null		Locale: Canada(English)	
null		Sort flags: Case-insensitive Column width: Unknown	
null	Describe	Status: Valid	
null	null	G Domain (20 of 551 members)	
null	13	li Null	
Sidney Lumet	null	Abel Ferrara	
null	97	Adam McKay	
Sam Firstenberg	null	Adam Shankman Adrian Lyne	
null	95	e	
Michael Polish	77	Ex Load Copy	

Figure 6

#### Question 5: How do I find all movies that contain specific strings in their names?

5.1 Create a group of all movies that contain "star" in their names. (Figure 7)

Hint: Triangle Click-> create group-> enter search criteria-> "Find ALL"-> Group ->Change group name (optional) -> Include Others(optional) -> Change Field name(optional)

Abc  movies.csv Title	Rename Copy Values Hide	ross	
Following		44,705	
Big Things	Aliases	0	
Hollywood Shuffle	Create Calculated Field	5,228,617	Create Group [Title]
Over the Hill to the Poor	Create Group	.000.000	Field Name: Title (group) 1 Groups: Add to:
Sweet Sweetback's Baa	Split Custom Split	,200,000	10,000 B.C. 102 Dalmatians 10th & Wolf
20,000 Leagues Under t		.000,000	11:14 12 Angry Men
The Beast from 20,000 F		,000,000	12 Rounds 13 Going On 30 1408
I Married a Strange Person	Describe	203,134	15 Minutes 16 Blocks 16 to Life
First Love, Last Rites	10,876	10,876	1776 1941 2 Fast 2 Furious
Let's Talk About Sex	373,615	373,615	2 For the Money Group Rename Ungroup V Show Add Loci
12 Angry Men	0	0	Indude 'Other' Indud
51 Birch Street	84,689	84,689	Find members
American Ninja 2: The C	4,000,000	4,000,000	Contains   Range: (All)
42nd Street	2,300,000	2,300,000	Find All Find Nex
Twin Falls Idaho	985,341	1,027,228	Reset OK Cancel Appl

Figure 7

### PART 2: Creating Simple Visualizations with Tableau

After doing basic data manipulations, we are going to create some simple visualizations. Click on the "Sheet1" tab and you will see a layout similar to Figure 8.

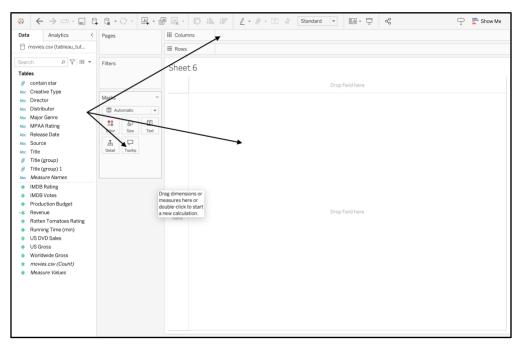


Figure 8: the arrows are indicating that you can drag and drop any of the measures or dimensions into any of the 3 areas shown.

The side bar has two important blocks, measures and dimensions.

Dimensions are values that determine the **level of detail at which measures are aggregated**. Basically, you will have a "mark" for each dimension (e.g., states in the US). You can think of them as slicing the

measures or creating groups into which the measures fit. The combination of dimensions used in the view defines the view's basic level of detail.

Measures are **values that are aggregated**. They can still be summed, averaged, counted, or they can have a minimum or maximum.

#### Question 6: Who is the most successful director?

6.1 Drag "Director" to the row, drag "Revenue" to the column. The "Show Me" menu let you create different types of charts (Figure 9).

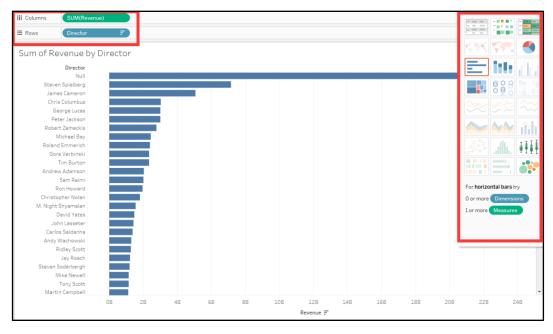


Figure 9

6.2 Which director has the highest sum of revenue? (Try to create a different visualization and sort the revenue) (Figure 10)

6.3 Filter Null Values in the "Director" field by right-clicking on the "Null" value and excluding the values. (notice that a new filter is created)

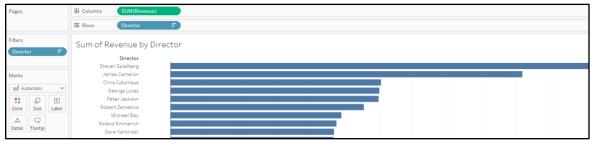


Figure 10

6.4 Drag "Revenue" to column and convert it to the measure of average (i.e.) (Figure 11).

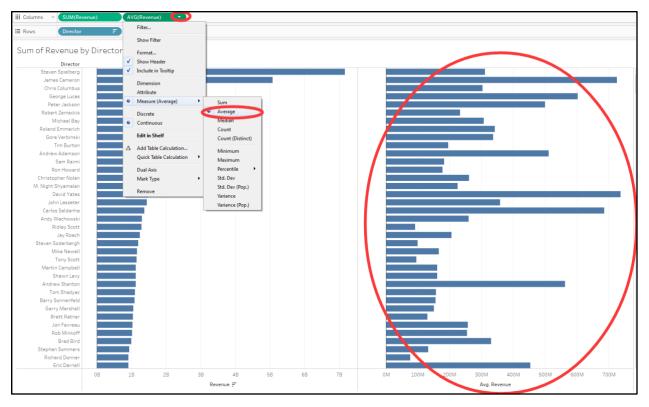


Figure 11

#### 6.5 Add annotations (Figure 11.1)

Director			
David Yates			
James Cameron	Select All		
Carlos Saldanha	View Data		
David Slade	 Сору	•	
George Lucas		ŕ	
Andrew Stanton	Format		
Pete Docter	Mark Label	•	
Richard Marquand	Annotate	•	Mark
Andrew Adamson	Trend Lines	•	Point
Peter Jackson	Forecast	►.	Area
Eric Darnell	Drop Lines	•	Ared

Figure 11. 1: Right click the field/object that you want to annotate

6.6 Edit the color of the bar by clicking on the "Color" option in the "Mark" shelf.

6.7 Edit the Sheet1 title to be "Sum and Average of Revenue by Director".

#### Question 7: Are there any directors who have never directed a profitable movie?

7.1 Create a new worksheet and name it as "Revenue as Dimension".

7.2 Drag "Director" to the Rows. Drag "Revenue" to Columns, and convert it to Dimension (Figure 12).

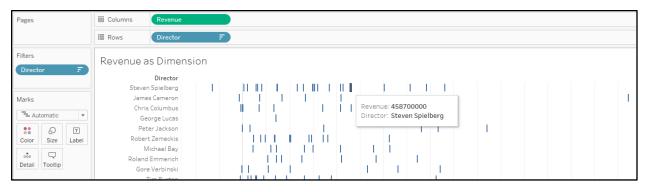


Figure 12

7.3 Add "Title" to the "Mark" shelf as tooltip so that each revenue mark is associated with the movie title (Figure 13)

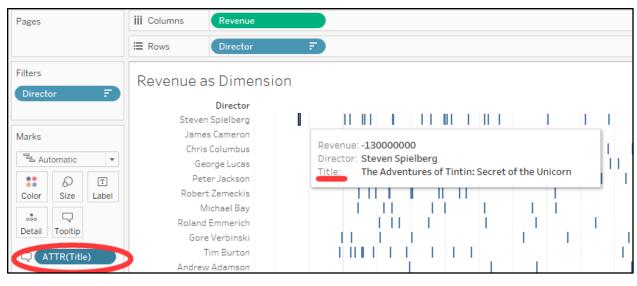


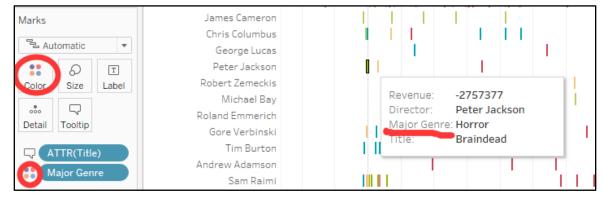
Figure 13

7.4 Add a constant line to indicate movies that have negative revenue (Hint: Analytics-> Summarize->Constant Line-> Value=0) (Figure 14)



Figure 14

#### 7.5 Drag "Major Genre" to "Marks" shelf and encode it using color. (Figure 15)





7.6 Are there directors who have never directed profitable movies? Who are they?

Hints:

- 1. You can use filters to complete this task.
- 2. Never directed profitable movie =all movies directed have revenues less than zero

MAX(Revenue) -	Sheet 6	Filter [Max. Revenue]
Edit Filter	Director Abel Ferrara	
Show Filter Clear Filter Apply to Worksheets	Akira Kurosawa Alan Rudolph Albert Brooks	Range of values
Dimension Attribute	ndrew Bergman Andy Garcia	
✓ Measure (Maximum) Add Table Calculation Quick Table Calculation Remove	Sum Average Median Count Count (Distinct)	-50,000,000 0 -50,000,000 2,530,891,499
	Bil Minimum ✓ Maximum Percentile ► Br Std.Dev C Std.Dev (Pop.) Variance Variance (Pop.)	Show: Only Relevant Values Include Null Values Reset Apply Cancel OK

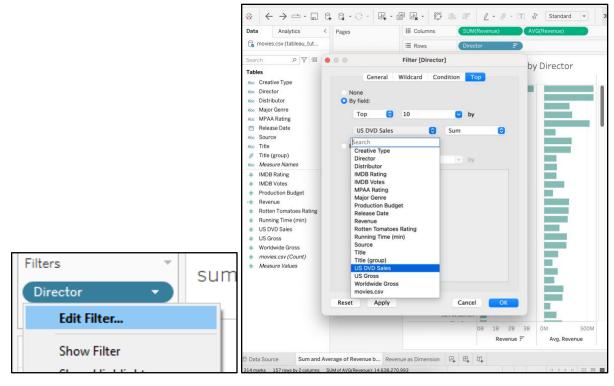
#### Question 8: Can I get more information about a specific genre?

8.1 Go back to the "Sum and Avg of Revenue" worksheet. Drag "Major Genre" to the "Filter" shelf and display only movies from the genre "Action". (Figure 16)

Filters		Filter [Major Genre]	×
Director 📻	sum	General Wildcard Condition Top	
	Director		
	David Yates	Select from list O Custom value list O Use all	=
Marks	James Cameron		
V All DOD	Carlos Saldanha	Enter search text	
✓ All 100	David Slade		^
∧ SUM(Reve 000	George Lucas	Action	
In Automatic 🔻	Andrew Stanton	Black Comedy	
	Pete Docter		
5 D I	Richard Marquand	Concert/Performance	
Color Size Label	Andrew Adamson	Documentary	
	Peter Jackson	Drama	
Detail Tooltip	Eric Darnell	Horror	

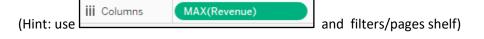


8.2 Find Directors who have the top 10 DVD sales by adjusting the director filter (Figure 17).





8.3 Which director has directed the most profitable Comedy movie?



#### Question 9: When is the best time to release a movie?

Create a visualization of revenue by quarters.

Hint: Click on the plus sign next to the YEAR (Figure 18)



Figure 19

### PART 3: Putting it all together

#### Question 10: How do I compare two visualizations in the same view?

10.1 Create a dashboard to host both worksheets that we have created (e.g. "Sum and Avg of Rev" and "Rev as Dimension") (Figure 19).

10.2 Select "Action" from the legend so that only action movies are displayed in the (Revenue as Dimension visualization) (Figure 19).

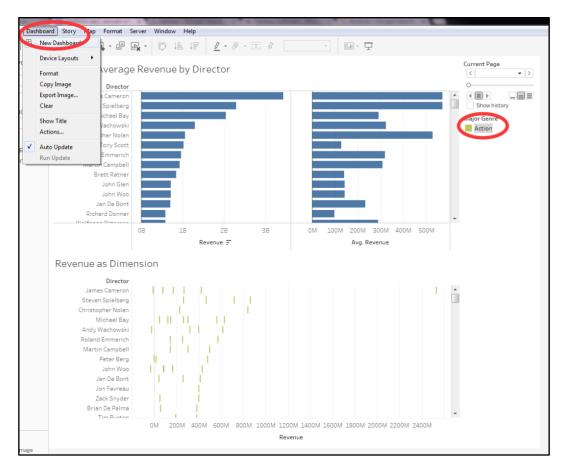


Figure 19

10.3 Add a highlight action: "Dashboard"-> "Action"-> "Add filter". In the pop-up window, add a "Hover" action for the field of "Director". (Figure 20).

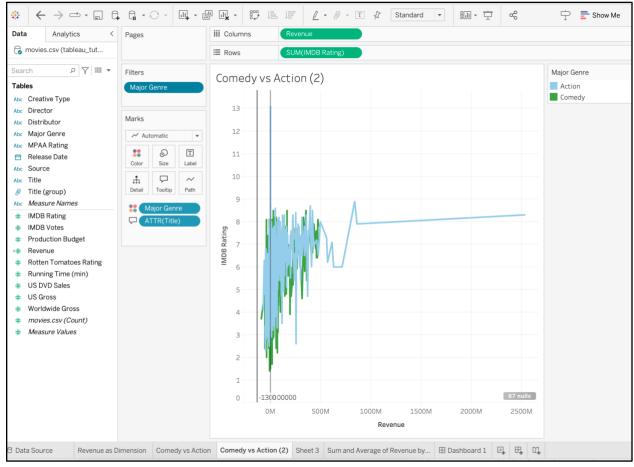
Dashboard Story Analysi	is Map Format		Server 1	Window Help 题	4					_	_ onlow me
New Dashboard		🛞 Table	au - practice-	tutorial	Rever	ue a	s Dimension	Sum and Averag	e of Rev	venue by Director 🛱	Major Genre
Device Layouts	> (	e 🖡 - 📪 🛍 🖉	2 - 0	- T 🖉 Standard		Di	Add Highlight Action		×	•	Action
Show Grid Grid Options	n	nension		Sum and Averag		Abel F	Name				Adventure Black Come
Format Copy Selected Dashboard Item		I In man		Director 🗐	Al	Adriai ira Kure	Highlight1	Insert	• 10	00M 0M 200M 400M	Concert/Pe
Copy Image				John Lasseter Gore Verbinski		Actio	Source Sheets		>	Max. Revenue	Drama Horror
Export Image Clear				OM 5		Action	Dashboard 1	Run action on	s, and		Musical
Show Title Actions	¢ # D				, An	the w Show	Revenue as Dimension     Sum and Average of Revenue by Director	Hover     Select     Menu			Romantic C Thriller/Sus Western
✓ Auto Update Run Update			1		A	Narr	Target Sheets				
Add Phone Layouts to Existing E Add Phone Layouts to New Dasi					Andr		Dashboard 1				
	Andrew Niccol Andrzej Bartkowiak Andy Garcia Andy Tennant				Ar		Revenue as Dimension Sum and Average of Revenue by Director				
	Andy Wachowski Ang Lee Angela Robinson				An		Target Highlighting				
00 Horizontal Container	Antonio Banderas		- Y'		Ari S.	Add	Dates and Times     Selected Fields		0		
B Vertical Container A Text È Extension © Pulse Metric	Ari S. Kirschenbaum Arthur Hiller Atom Egoyan Barbet Schroeder					ora stre arry Lev	ATTR(Title)  Director  Major Genre				
Data Story     Image     Blank	Barbra Streisand Barry Levinson Barry Sonnenfeld				Ban	y Sonn Ben A	Revenue				
왕 Workflow ④ Web Page Tiled Floating	Ben Affleck Betty Thomas Bill Paxton				]	Betty Th Bill F Bille A	ugust  -130000000 0	Cancel OK			
Show dashboard title	Bille August	-130000000	0		s Dimension	Sheet 3	Sum and Average of Revenue by   Dashboard 1	1 R. H. M.			-
8 Data Source Revenue as D	Dimension Sheet 3 Sum	and Average of Revenue by	Dashboard			-	Consideration of the second of				► H

Figure 20

*Question 11: Is it better to invest in action movies or in comedy movies? Which director would you invest in?* 

Data	Analytics <	Pages	III Columns Revenue	
0	novies.csv (tableau_tut		E Rows Director	
eard	:h _ ₽ 🕅 ¥	Filters	Comedy vs Action	Major Genre
able	95	Major Genre	Connedy vs Action	Action
bo	Creative Type		Director	Comedy
	Director		Null	
bo	Distributor	Marks	Adam McKay	
be	Major Genre	🖏 Automatic 💌	Adam Shankman	
	MPAA Rating		Akira Kurosawa	
	Release Date	1 😞 🗉	Alan Alda	
	Source	Color Size Label	Alan Rudolph	
	Title		Albert Brooks	
	Title (group)	Detail Tooltip	Albert Hughes	
	Measure Names		Alex Proyas	
	IMDB Rating	Sector Major Genre	Amy Heckerling	
	IMDB Votes	ATTR(Title)	Andrei Konchalovsky Andrew Bergman	
	Production Budget		Andrew Bergman Andrew Davis	
	Revenue		Andrew Fleming	
	Rotten Tomatoes Rating		Andrew Niccol	
	Rotten Tomatoes Rating Running Time (min)		Andrew Niccol	
	US DVD Sales		Andrzej Bartkowiak	
	US DVD Sales US Gross		Andy Fickman	
	Worldwide Gross		Andy Tennant	
			Andy Wachowski	
	movies.csv (Count) Measure Values		Anglee	
H¢.	measure values		Angela Robinson Revenue: 611576929	
			Antoine Fuqua Director: Andy Wachowski Major Genre: Action	
			Arthur Hiller Title: The Matrix Reloaded	
			Barry Levinson	
			Barry Sonnenfeld	
			Ben Stiller -13000000 2 nulls	
			0M 500M 1000M 1500M 2000M 2500M	
			Revenue	
		mension Comedy vs Acti	n Sheet 3 Sum and Average of Revenue by 🖽 Dashboard 1 🖳 🖽 🖽	

### Question 12: What's the relationship between IMDB ratings and the revenue? Are highlyrated movies more profitable?



#### Question 13: I am going to invest in ......

13.1 Make a story by clicking on the "add story" button	$\Box$	
13.1 Make a story by clicking on the "add story" button		(figure 21)

	Story	Layout	\$	
	New story	point		
	Blank	Duplicate		
13.2 Annotate the story and add new story points.			_	(figure 22)

13.3 Who would you invest? Create some visualizations to justify your decision.

### PART 4: Exporting your visualizations

Interactive Tableau workbooks can only be uploaded to Tableau online, Tableau Server and Tableau servers. However, you can embed these visualizations to your own website: see the instruction here: <a href="https://onlinehelp.tableau.com/current/api/js\_api/en-us/JavaScriptAPI/js\_api.htm">https://onlinehelp.tableau.com/current/api/js\_api/en-us/JavaScriptAPI/js\_api.htm</a>

You can export to image and pdf also.

For image export, you can only export worksheets: (figure 23)

_											
Tab	leau -	Boo	k1								
0	)ata	Wor	rksheet	Dashboard	Story	An	alysis	Map	Form	nat S	erv
	÷	<u>.</u>	New W	/orksheet	Ctrl+	М	4 -		<b>:</b>	₽,2	ţ
ata			Сору			►			100	Colu	m
			Export			•		lmage.			E
6 r	novie		Clear			×		Data			Ŀ
me	ensic		Action	s				Crossta	ab to Ex	cel	F
	Crea Direc		Tooltip						S	um	
oc	Distr	$\checkmark$	Show T	litle							
oc	Majo		Show (	Caption							
oc	MPA		Show S	Summary			L				
9	Relea		Show (	Cards		►		00	]		
с	Sour		Show V	/iew Toolbar		•	tic	Ŧ			
9	star į		Descrik	oe Sheet	Ctrl+	c					
oc	Title					-	P	T			
с	Mea:		Duplica	ate as Crosstab	,		ze	Label			
			Auto U	pdates		►	2				
eas	sure		Run Up	odate		Þ	ltip				

Figure 23: export by clicking worksheet-> export->image

For pdf export (which you can edit on editors such as illustrator), go to file-> print to pdf (figure 24)

Dashboard Layout <	Revenue as Dimensio	-	Cum and Au		nue hu Director
Default	Revenue as Dimension	a	Sum and Av	erage of Rever	nue by Director
Phone	Director		Director =		
	Null		Shawn Levy		
Device Preview	Abel Ferrara		John Lasseter		
		Print			
Size				4 500M 1000M	/ OM 200M 400N
Desktop Browser (1000 x 8 *				Revenue =	Max. Revenue
	Printer	🧂 Library - Printer and	Scanner (Library)		
Sheets	The printer is offline.				
Revenue as	Presets		Default Settings 🗘		
Comedy vs Action					
Cornedy vs Action					
Sheet 3	Copies		1 0		
Sum and Average o					
I Sheet 6	Print in Color				
	Double-sided		On 💲		
	<ul> <li>Tableau</li> </ul>				
Dbjects	Print Ran	nge: 🗿 Entire workbook			
D Horizontal Container		Active sheet			
Vertical Container		<ul> <li>Selected sheets</li> </ul>			
A Text					
> Extension	Show	selections			
Pulse Metric					
1) Data Story	Madia & Quality				
년 Image	> MORIS & LIUSIN				
Blank	? PDF v		Cancel Print		
詞 Workflow					
Web Page	Betty Thomas				
	Bill Paxton				
Tiled Floating	Bille August -1300000		0		

Figure 24

Note: Make sure that you check out the tableau public gallery <u>https://public.tableau.com/en-us/s/gallery</u>