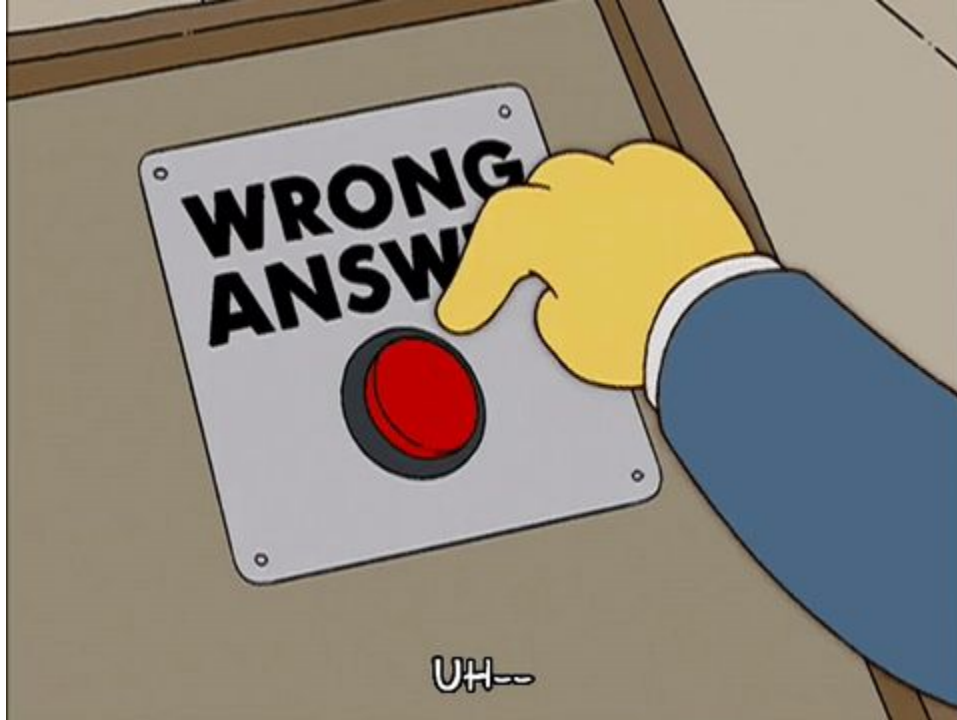


How to be RIGHT

About being wrong...

Prof. Ernesto Lee

Learn to be wrong... the right way...



Journey



Class Routine (Beginning)

- Ask yourself this question: What belief will I be wrong about today?
- Decide how you'll seek out information to disprove your belief.
- Prepare for blowback.



Class Routine (End)

1. Ask yourself this question: What was I proven wrong about today? Once again, document your answer.
2. Think about how you were challenged or proven wrong.
3. Reward yourself if you made it through the day and successfully disproved something you believed, or at least opened yourself for future questioning.



Transformation

- I don't want you to learn Data Analytics
- I want you to learn HOW to teach YOURSELF Data Analytics



If I won the lottery, the first thing I would do is...



JESSE



**I CAN
PREDICT THE
STOCK
MARKET. MY
ALGORITHMS
BEAT THE
MARKET!!**



300,000 PEOPLE



**100K: THIS STOCK
PRICE WILL RISE**



**100K: THIS STOCK
PRICE WILL FALL**



**100K: THIS STOCK
PRICE WILL STAY
THE SAME**

100,000 PEOPLE
(week 2)

**33K: THIS STOCK
PRICE WILL RISE**

33,000 PEOPLE

**11K: THIS STOCK
PRICE WILL RISE**

**11,000
PEOPLE
(PAY)**

**3.5K: THIS
STOCK
PRICE WILL
RISE**

**33K: THIS STOCK
PRICE WILL FALL**

**11K: THIS STOCK
PRICE WILL FALL**

**3.5K: THIS
STOCK
PRICE WILL
FALL**

**33K: THIS STOCK
PRICE WILL STAY
THE SAME**

**11K: THIS STOCK
PRICE WILL STAY
THE SAME**

**3.5K: THIS
STOCK
PRICE WILL
STAY THE
SAME**

Essence Document

1-sentence course summary

- “Create predictive models with Linear Algebra and improve them with Calculus.”
 - Linear Algebra efficiently describes complex operations using simple steps.
 - Calculus can optimize any model with derivatives (gradient). Minimize error.

1-sentence core concepts

- Linear Algebra: spreadsheets for your equations. We "pour" data through various operations.
- Natural log: time needed to grow. Helps normalize widely varying numbers.

Project ideas

- Correlation between words in title and popularity / comments / traffic

AHA Moments

- When the lightbulb clicks
- Make sure to highlight KEY INSIGHTS

Huh?

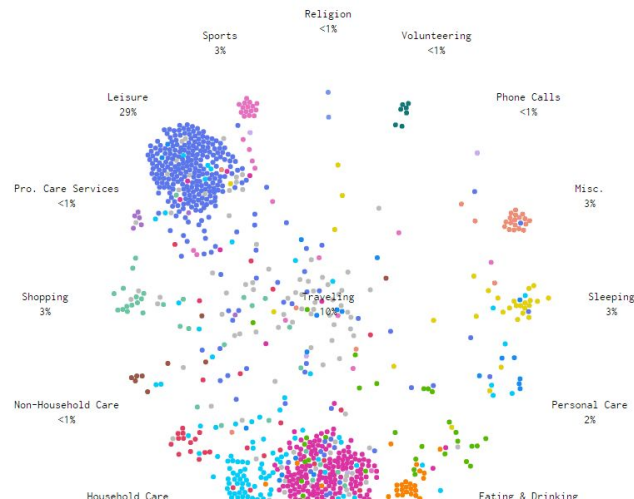
- Questions you still have unanswered

Data Visualization

- Showing basic comparisons and relationships between variables
- Building data distributions using histograms
- Representing geographical distribution of data in maps
- Showing data that changes over time

3:36pm

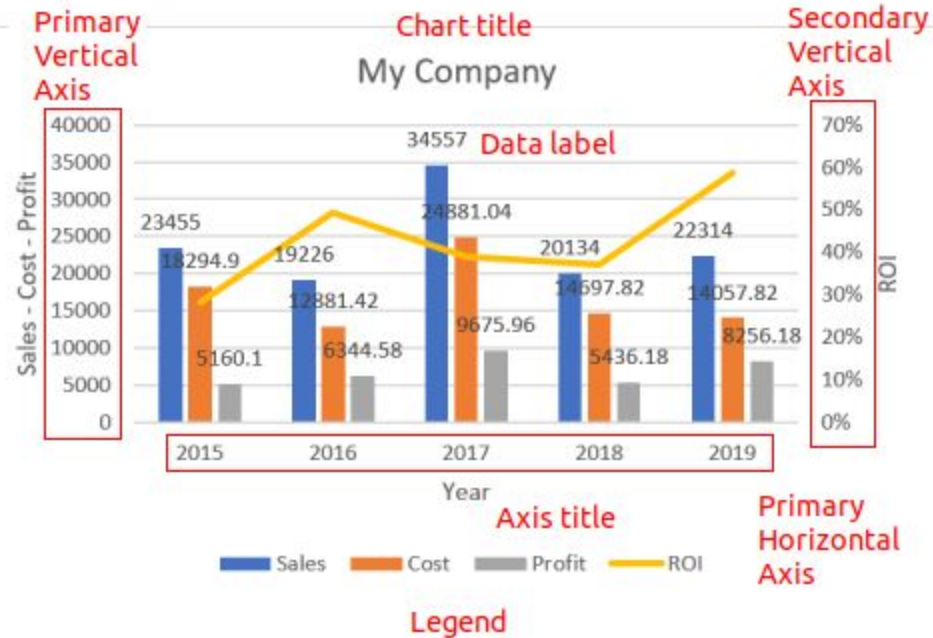
Coffee break? Again, at the top of the hour, you see a shift in activity.



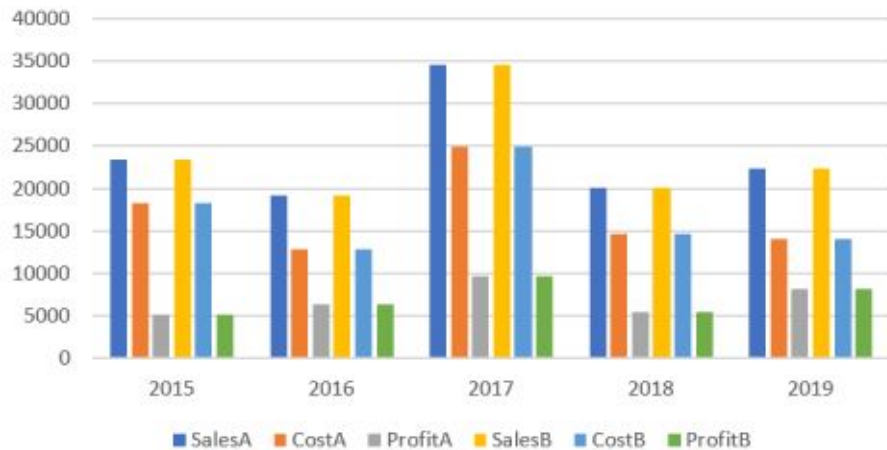
Showing basic comparisons and relationships between variables

1. Tell the story of your data and help decision makers with their job.
2. Predict the future evolution of some variable(s).
3. Find hidden trends and patterns in the data.
4. Find outliers, that is, anomalies in the data.
5. Understand the distribution, composition, and relationships.
6. Build groups and categories.

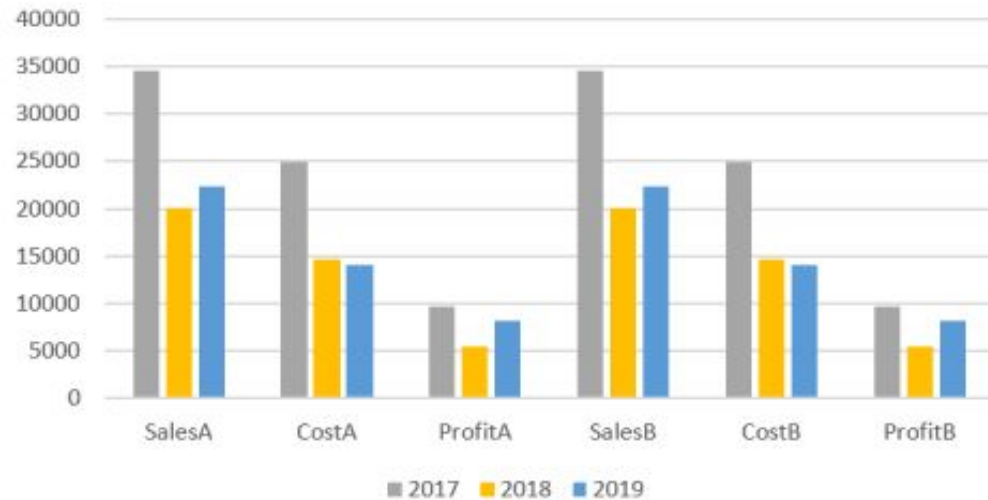
Year	SalesA	CostA	ProfitA	SalesB	CostB	ProfitB
2015	23455	9	1	23455	9	1
2016	19226	42	58	19226	42	58
2017	34557	04	96	34557	04	96
2018	20134	82	18			
2019	22314	82	18			



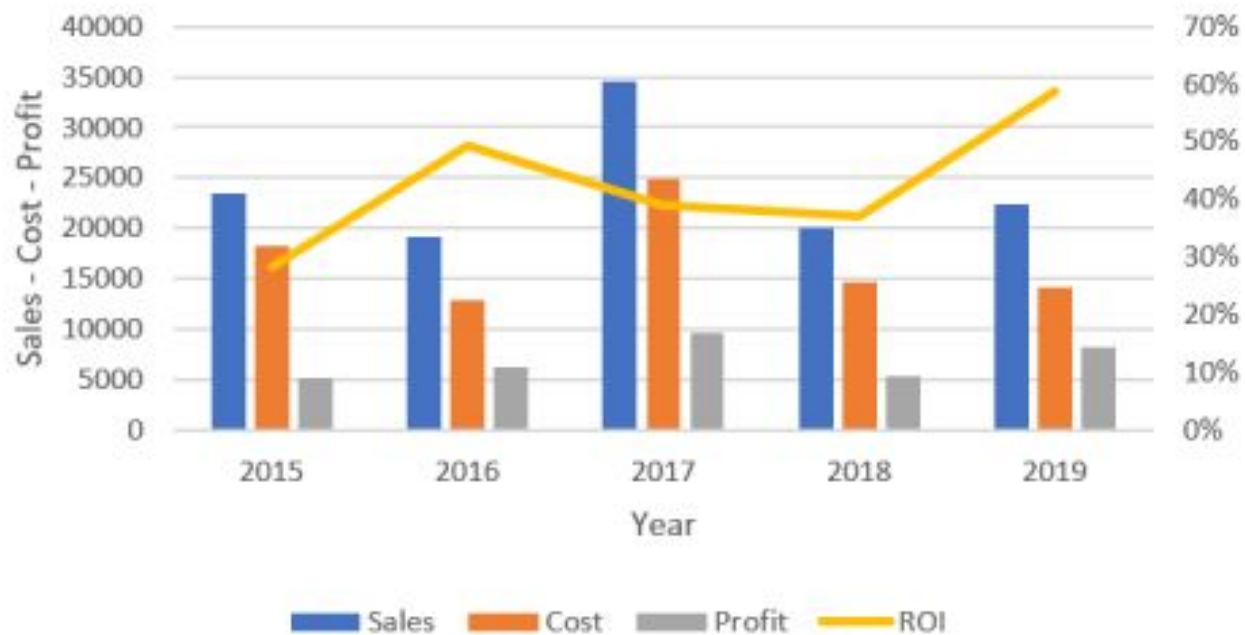
My Company



My Company



My Company

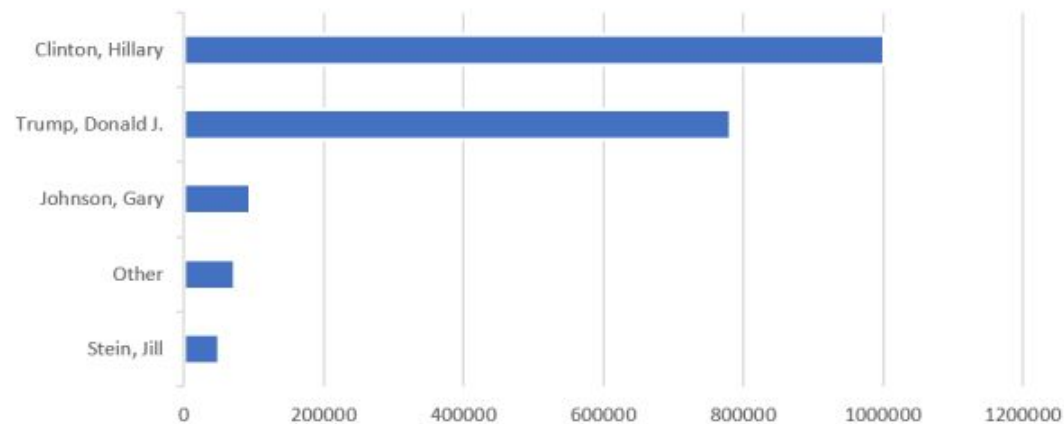


My Company

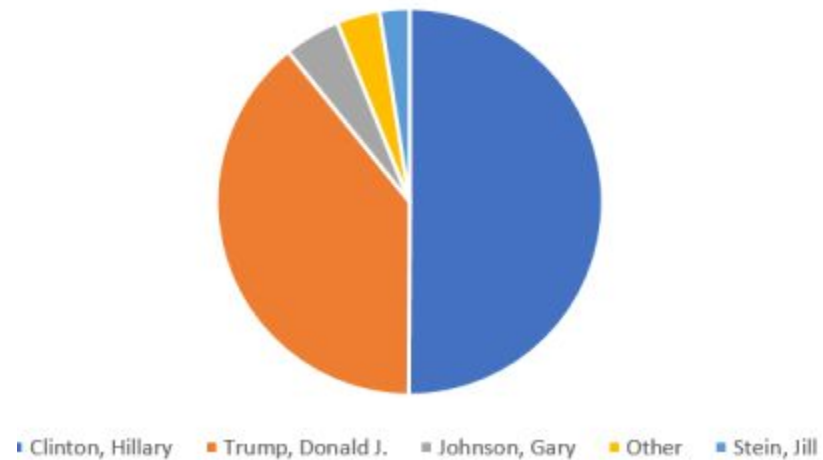


Candidate	Votes
Clinton, Hillary	1,002,106
Trump, Donald J.	782,403
Johnson, Gary	94,231
Other	72,594
Stein, Jill	50,002

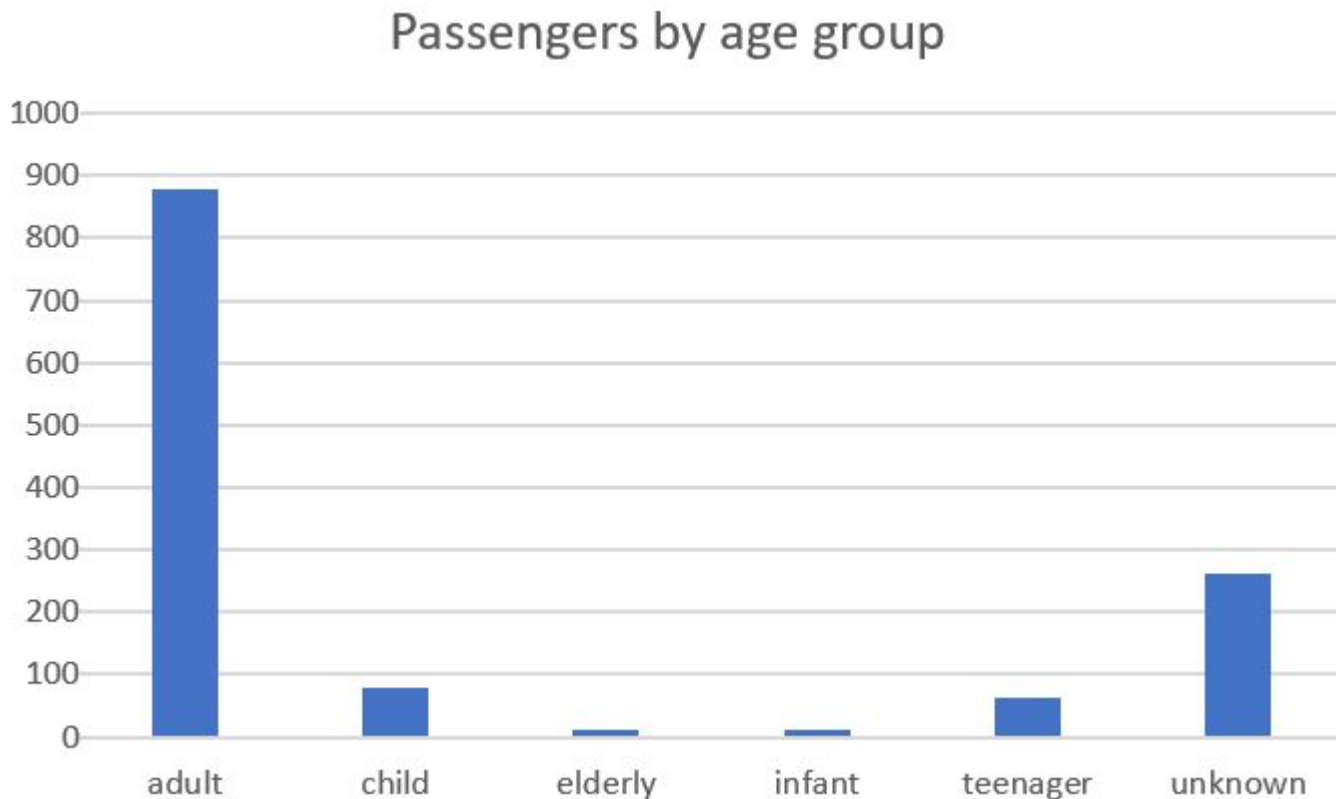
US President 2016 election - Oregon

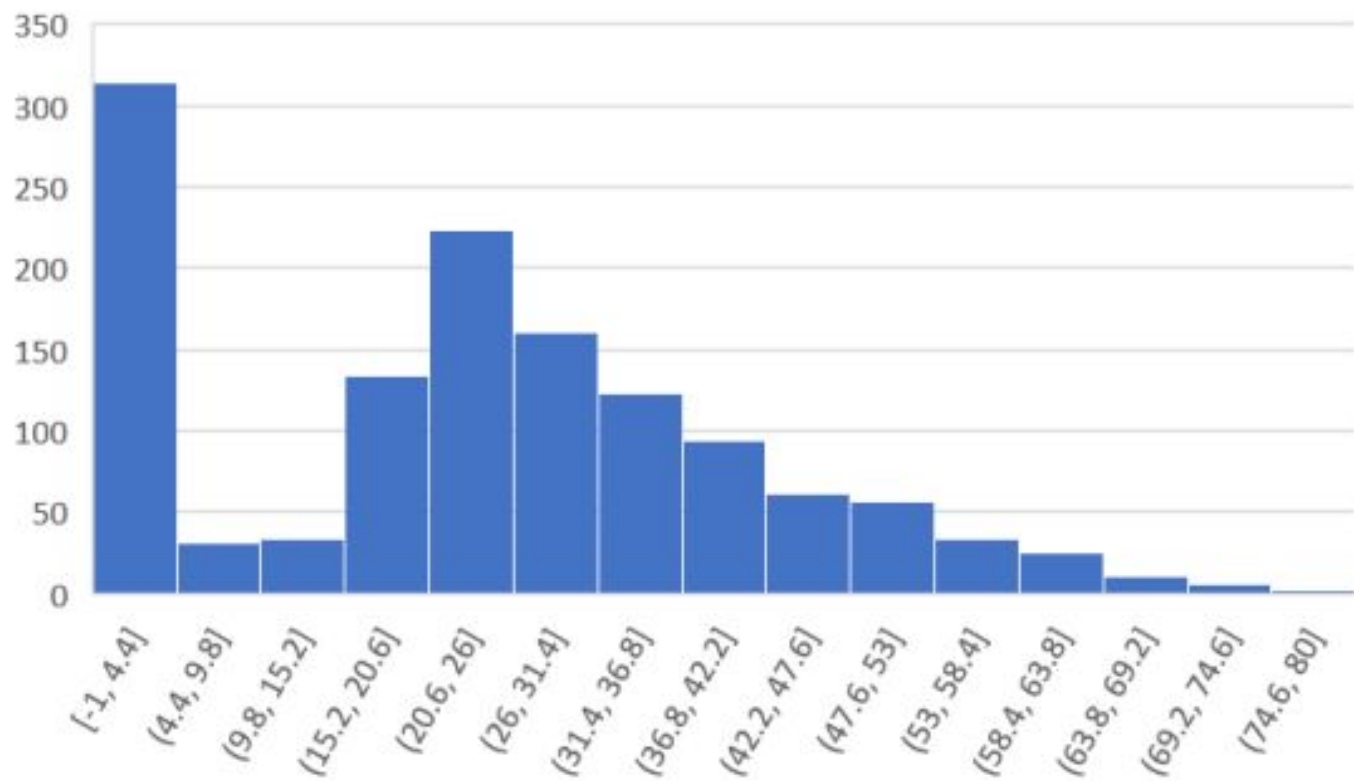


US President 2016 election - Oregon



Importance of Pre-Processing





	A	B	C	D	E	F	G	H	I	J	K	L	M
1	year	state	state_po	state_fips	state_cen	state_ic	office	candidate	party	writein	candidatevotes	totalvotes	version
2	1976	Alabama	AL	1	63	41	US President	Carter, Jimmy	democrat	FALSE	659170	1182850	20171015
3	1976	Alabama	AL	1	63	41	US President	Ford, Gerald	republican	FALSE	504070	1182850	20171015
4	1976	Alabama	AL	1	63	41	US President	Maddox, Lester	american independent party	FALSE	9198	1182850	20171015
5	1976	Alabama	AL	1	63	41	US President	Bubar, Benjamin "Ben"	prohibition	FALSE	6669	1182850	20171015
6	1976	Alabama	AL	1	63	41	US President	Hall, Gus	communist party use	FALSE	1954	1182850	20171015
7	1976	Alabama	AL	1	63	41	US President	Macbride, Roger	libertarian	FALSE	1481	1182850	20171015
8	1976	Alabama	AL	1	63	41	US President			TRUE	308	1182850	20171015
9	1976	Alaska	AK	2	94	81	US President	Ford, Gerald	republican	FALSE	71555	123574	20171015
10	1976	Alaska	AK	2	94	81	US President	Carter, Jimmy	democrat	FALSE	44058	123574	20171015
11	1976	Alaska	AK	2	94	81	US President	Macbride, Roger	libertarian	FALSE	6785	123574	20171015
12	1976	Alaska	AK	2	94	81	US President			TRUE	1176	123574	20171015
13	1976	Arizona	AZ	4	86	61	US President	Ford, Gerald	republican	FALSE	418642	742719	20171015
14	1976	Arizona	AZ	4	86	61	US President	Carter, Jimmy	democrat	FALSE	295602	742719	20171015

File Home Transform Add Column View

Close & Load Close Refresh Preview Query Properties Advanced Editor Manage Choose Columns Remove Columns Manage Columns Keep Rows Remove Rows Reduce Rows Sort Split Column Group By Data Type: Any Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Combine Parameters Data source settings Data Sources New Source Recent Sources New Query

Queries

fx = Excel.CurrentWorkbook()[[Name="Table1"]][Content]

	year	ABC 123 state	ABC 123 state_po	ABC 123 state_fips	ABC 123 state_cen	ABC 123 state_ic	ABC 123 office	ABC 123 candidate	ABC 123
1	1976	Alabama	AL		1	63	41 US President	Carter, Jimmy	
2	1976	Alabama	AL		1	63	41 US President	Ford, Gerald	
3	1976	Alabama	AL		1	63	41 US President	Maddox, Lester	
4	1976	Alabama	AL		1	63	41 US President	Bubar, Benjamin "Ben"	
5	1976	Alabama	AL		1	63	41 US President	Hall, Gus	
6	1976	Alabama	AL		1	63	41 US President	Macbride, Roger	
7	1976	Alabama	AL		1	63	41 US President		null



Group By

☐ Basic ☒ Advanced

Specify the columns to group by and one or more outputs.

Group by

state ▼

Add grouping

New column name

Winner

Votes

Add aggregation

Operation

Sum ▼

Max ▼

Column

party ▼

candidatevotes ▼

OK

Cancel

Queries

✕ ✓ π = Table.Group("#Changed Type", {"state"}, {{ "Winner", each List.Sum([party]), type text}, {"Votes", each List.Max

	state	Winner	Votes
1	Alabama	Error	659170
2	Alaska	Error	71555
3	Arizona	Error	418642
4	Arkansas	Error	498604
5	California	Error	3882244
6	Colorado	Error	584278
7	Connecticut	Error	719261
8	Delaware	Error	122461
9	District of Columbia	Error	137818
10	Florida	Error	1636000
11	Georgia	Error	979409

File Home Transform Add Column View

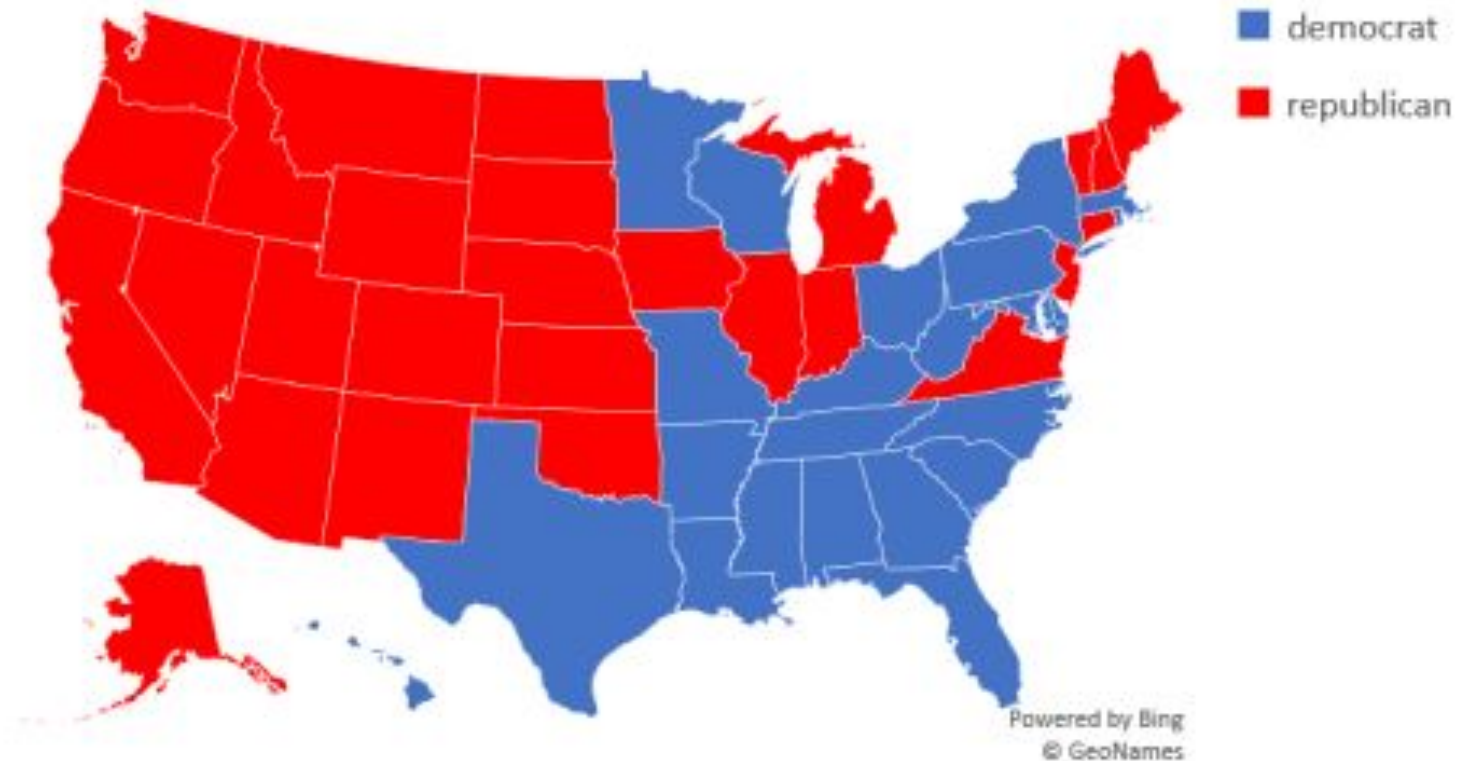
Close & Load Close Refresh Preview Query Properties Advanced Editor Manage Choose Columns Remove Columns Manage Columns Keep Rows Remove Rows Reduce Rows Sort Split Column Group By Data Type: Text Use First Row as Headers Replace Values Transform Merge Queries Append Queries Combine Files Combine Parameters Data source settings Data Sources New S Recent New 4

Queries

✕ ✓ fx = Table.Group("#Changed Type", {"state"}, {"Winner", each List.First([party]), type text}, {"Votes", each List.Max

	state	Winner	Votes
1	Alabama	democrat	659170
2	Alaska	republican	71555
3	Arizona	republican	418642
4	Arkansas	democrat	498604
5	California	republican	3882244
6	Colorado	republican	584278
7	Connecticut	republican	719261
8	Delaware	democrat	122461

1976 US President election results



	A	B	C	D
1	Long	Lat	Station	Line
2	-58.39892759	-34.63575018	CASEROS	H
3	-58.40096956	-34.62937566	INCLAN - MEZQUITA AL AHMAD	H
4	-58.40232273	-34.62309232	HUMBERTO 1°	H
5	-58.40473172	-34.61524215	VENEZUELA	H
6	-58.40603638	-34.60893524	ONCE - 30 DE DICIEMBRE	H
7	-58.38057434	-34.6042452	9 DE JULIO	D
8	-58.39792376	-34.59975708	FACULTAD DE MEDICINA	D
9	-58.38514236	-34.60158717	TRIBUNALES - TEATRO COLÓN	D
10	-58.40716132	-34.59162784	AGÜERO	D
11	-58.41595542	-34.58515594	R.SCALABRINI ORTIZ	D
12	-58.42119601	-34.58141119	PLAZA ITALIA	D
13	-58.42571144	-34.57842202	PALERMO	D
14	-58.37401822	-34.59119381	RETIRO	C
15	-58.37815578	-34.60176992	LAVALLE	C
16	-58.37952998	-34.60484374	DIAGONAL NORTE	C
17	-58.38061072	-34.60898331	AV. DE MAYO	C
18	-58.38044447	-34.61261728	MORENO	C
19	-58.38017361	-34.6181256	INDEPENDENCIA	C
20	-58.38143443	-34.62761945	CONSTITUCION	C
21	-58.37507152	-34.60329729	FLORIDA	B
22	-58.38071485	-34.60363711	C. PELLEGRINI	B
23	-58.38729613	-34.60409355	URUGUAY	B
24	-58.39231424	-34.60441954	CALLAO - MAESTRO ALFREDO BRAVO	B
25	-58.39947426	-34.60464297	PASTEUR - AMIA	B
26	-58.40539944	-34.60458106	PUEYRREDON	B
27	-58.4117626	-34.60407952	CARLOS GARDEL	B

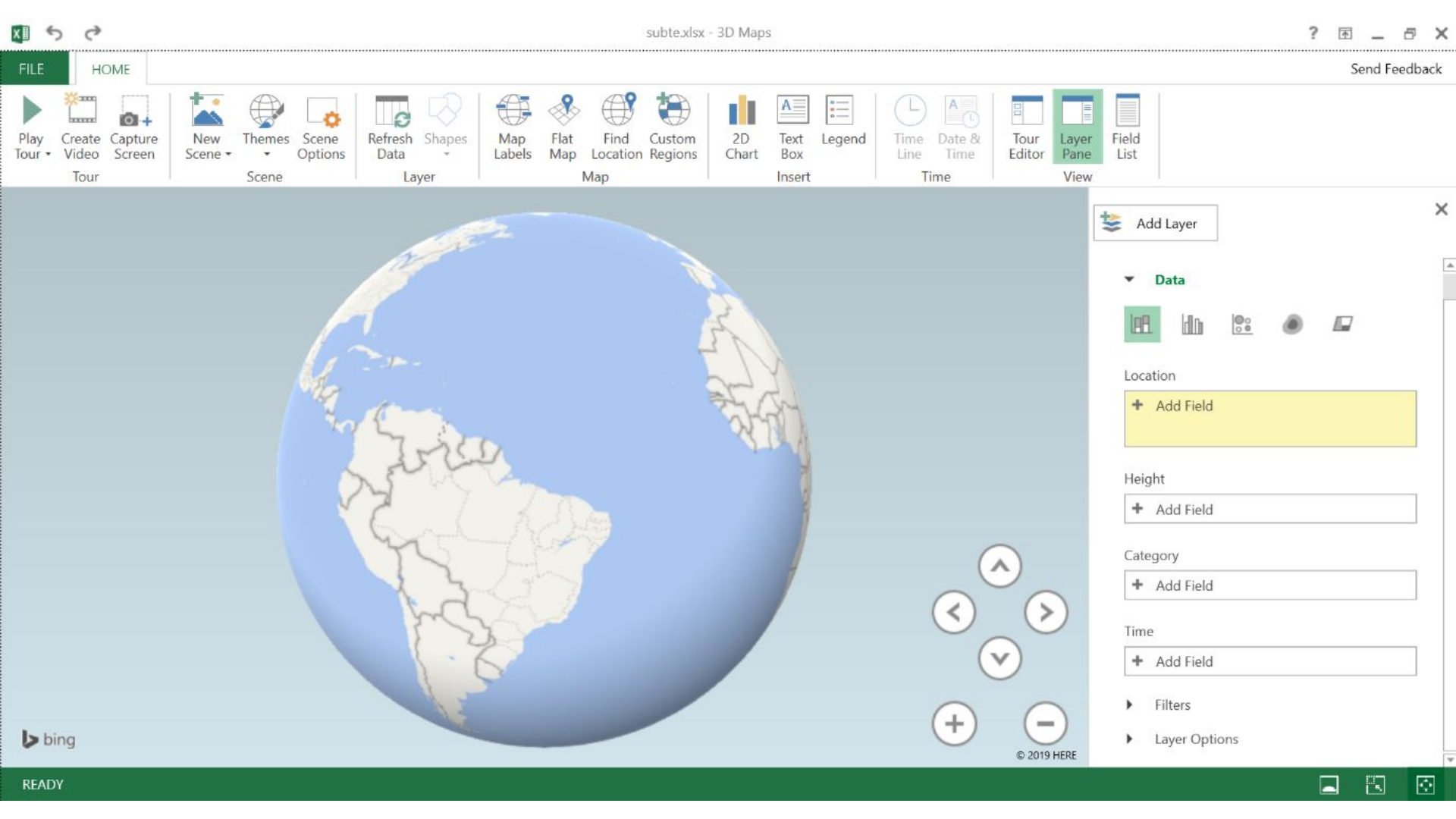
	File	Home	Insert	Page Layout	Formulas	Data	Review	View	Developer	Add-ins	Master Data	Inquire	Power Pivot	PDF-XChange	Team	Tell me what you want to do	Share				
	PivotTable	Recommended PivotTables	Table	Pictures	Online Pictures	Store	My Add-ins	Recommended Charts	Charts	Maps	PivotChart	3D Map	Line	Column	Win/Loss	Slicer	Timeline	Link	Text	Equation	Symbol
	Tables			Illustrations		Add-ins						Tours	Sparklines		Filters	Links					
64	-58.44466815	-34.57001231	OLLEROS																		
65	-58.4521256	-34.56621524	JOSE HERNANDEZ																		
66	-58.45648913	-34.56230909	JURAMENTO																		
67	-58.45502929	-34.58719785	FEDERICO LACROZE																		
68	-58.46165176	-34.64331216	PLAZA DE LOS VIRREYES - EVA PERON																		
69	-58.45789176	-34.64013735	VARELA																		
70	-58.37395581	-34.60780234	CATEDRAL																		
71	-58.3709685	-34.60881031	PLAZA DE MAYO																		
72	-58.46237841	-34.55564177	CONGRESO DE TUCUMAN																		

3D Map

See your geographic data on a 3D map, visualized over time.

Explore it for insights, animate changes over time and create a video.

[Tell me more](#)



FILE

HOME

Play Tour

Create Video

Capture Screen

New Scene

Themes

Scene Options

Refresh Data

Shapes

Map Labels

Flat Map

Find Location

Custom Regions

2D Chart

Text Box

Legend

Time Line

Date & Time

Tour Editor

Layer Pane

Field List

subte.xlsx - 3D Maps

?

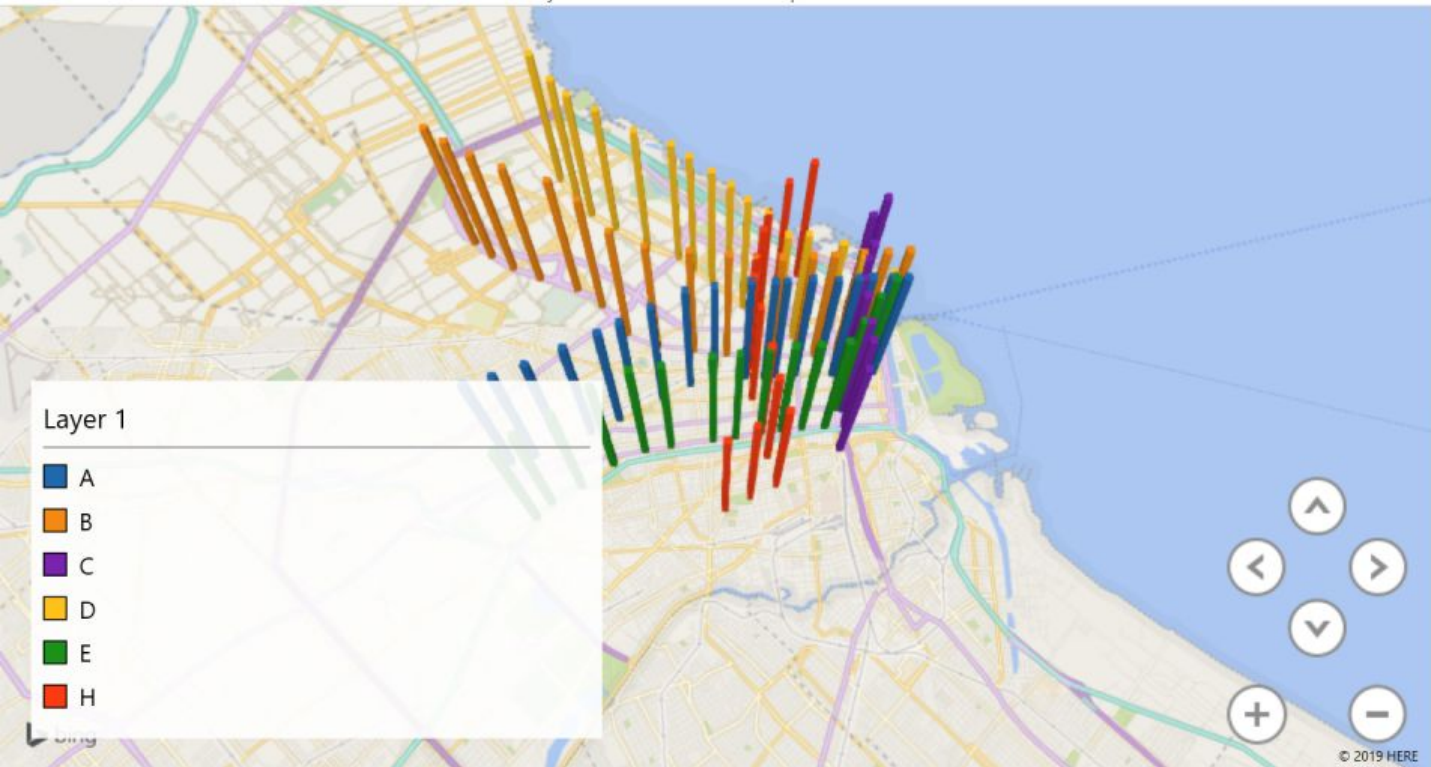
🔍

⌵

📄

✖

Send Feedback



+

Add Layer

✖

Location

☒ Lat

Latitude

✖

☐ Long

Longitude

✖

+

Add Field

Height

+

Add Field

Category

Line

✖

Time

+

Add Field

Filters

+

Add Filter

Layer Options

Height

100%



FILE

HOME

Play
TourCreate
VideoCapture
Screen

Tour

New
Scene

Themes

Scene
Options

Scene

Refresh
Data

Shapes

Layer

Map
LabelsFlat
Map

Map

Find
LocationCustom
Regions2D
ChartText
Box

Insert



Legend

Time
LineDate &
Time

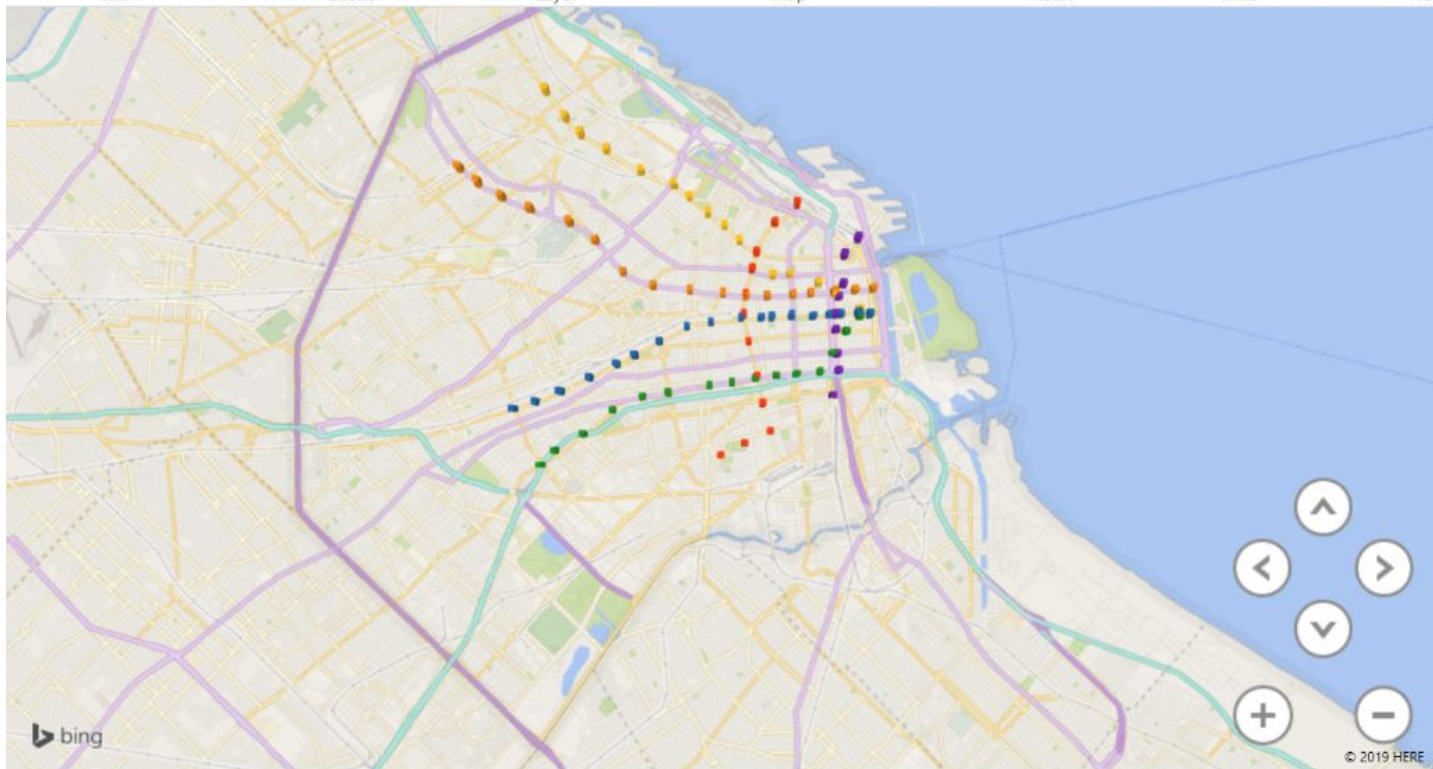
Time

Tour
EditorLayer
Pane

View

Field
List

Send Feedback



bing

© 2019 HERE



Add Layer

Filters

Layer Options

Height

 0%

Thickness

 50%☐ Lock current scale

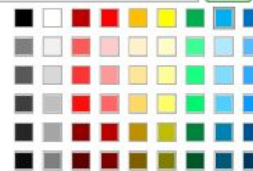
Opacity

 100%

Color

A

Show Values

☒ Zeroes☒ Negatives☐ Nulls

More Colors...

Reset layer colors

READY

FINISHED



Customize Data Card

CHOOSE DATA FIELDS FOR CUSTOM TOOLTIP

Lat

Long

Line

+ Add Field



Reset to Defaults

◀ TEMPLATE 1 ▶

Lat

Value ▶

Long

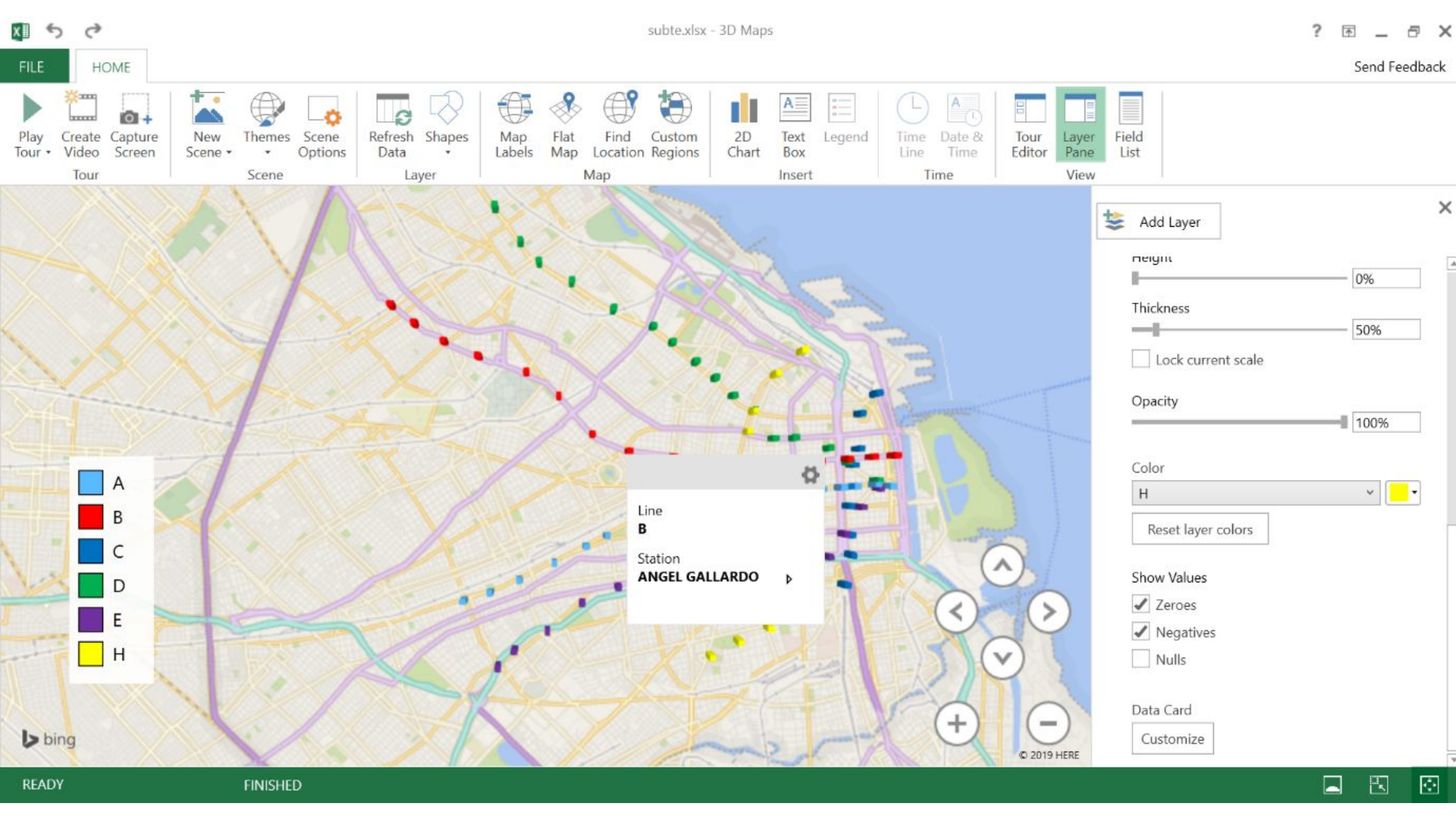
Value ▶

Line

Value ▶

OK

Cancel



Dynamic Visuals

Choose Columns

Choose the columns to keep

☒ (Select All Columns)
☒ year
☒ state
☐ state_po
☐ state_fips
☐ state_cen
☐ state_ic
☐ office
☐ candidate
☒ party
☐ writein
☒ candidatevotes
☒ totalvotes
☐ version
☐ notes

OK

Cancel

Queries

>

✕

✓

$\int x$

= Table.TransformColumnTypes(#"Added Custom",{{"Percentage", Percentage.Type}})

	1 ² 3 year	A ^B C state	A ^B C party	1 ² 3 candidatevotes	1 ² 3 totalvotes	% Percentage
1	1976	Alabama	democrat	659170	1182850	55.73 %
2	1976	Alabama	republican	504070	1182850	42.61 %
3	1976	Alabama	american independent party	9198	1182850	0.78 %
4	1976	Alabama	prohibition	6669	1182850	0.56 %
5	1976	Alabama	communist party use	1954	1182850	0.17 %
6	1976	Alabama	libertarian	1481	1182850	0.13 %
7	1976	Alabama	null	308	1182850	0.03 %
8	1976	Alaska	republican	71555	123574	57.90 %
9	1976	Alaska	democrat	44058	123574	35.65 %
10	1976	Alaska	libertarian	6785	123574	5.49 %
11	1976	Alaska	null	1176	123574	0.95 %
12	1976	Arizona	republican	418642	742719	56.37 %
13	1976	Arizona	democrat	295602	742719	39.80 %



Pivot Column

Use the names in column "party" to create new columns.

Values Column ⓘ

Percentage ▾

Advanced options

Aggregate Value Function ⓘ

Don't Aggregate ▾

[Learn more about Pivot Column](#)

Queries

×

✓

f_x

= Table.Pivot("#Filtered Rows", List.Distinct("#Filtered Rows"[party]), "party", "Percentage")

<div><div></div><div></div><div></div></div>	<div>1²3</div> year	<div>A^BC</div> state	<div>1²3</div> candidatevotes	<div>1²3</div> totalvotes	% republican	% democrat
1	1976	California	3742284	7803770	null	47.95 %
2	1976	California	3882244	7803770	49.75 %	null
3	1980	California	3082943	8582938	null	35.92 %
4	1980	California	4522994	8582938	52.70 %	null
5	1984	California	3922519	9505041	null	41.27 %
6	1984	California	5467009	9505041	57.52 %	null
7	1988	California	4702233	9887065	null	47.56 %
8	1988	California	5054917	9887065	51.13 %	null
9	1992	California	3630574	11131721	32.61 %	null
10	1992	California	5121325	11131721	null	46.01 %
11	1996	California	3828381	10019469	38.21 %	null
12	1996	California	5119835	10019469	null	51.10 %

	A	B	C	D
1	year ▼	Percentage Republican ▼	Percentage Democrat ▼	Other ▼
2	1976	49.75%	47.95%	2.30%
3	1980	52.70%	35.92%	11.38%
4	1984	57.52%	41.27%	1.22%
5	1988	51.13%	47.56%	1.31%
6	1992	32.61%	46.01%	21.38%
7	1996	38.21%	51.10%	10.69%
8	2000	41.65%	53.45%	4.90%
9	2004	44.36%	54.31%	1.34%
10	2008	36.95%	61.01%	2.03%
11	2012	37.12%	60.24%	2.64%
12	2016	31.62%	61.73%	6.66%

Presidential election Results - California

