Hands On: Cars Dataset

Fernando Birra
João Moura Pires

Notice

- **Author**
  - **João Moura Pires (jmp@fct.unl.pt)**
  - **Fernando Birra (fpb@fct.unl.pt)**

Some practical Information

- **You are 76 students**
  - **24 teams with 3 students and 1 team with 4 students**
  - Or
  - **24 teams with 3 students and 2 teams with 2 students**

- **All the students of a team should be placed in the same lab shift**

ID	Time	Actual
P1	Fri - 18	29
P2	Fri - 14	26
P3	Thu - 11	21

Table of Contents

- **Cars Dataset (Examples)**

Hands On: Cars Dataset

Hands On: Cars Dataset

- **Steps:**
  - **Get familiar with the cars dataset**
  - **Download the 2004 Cars and Trucks Data Set dataset**
  - **Use tableau to connect live to the excel file**

Metadata

Associated Metadata

```
NAME:  2004 New Car and Truck Data
TYPE:  Sample
SIZE:  428 observations, 19 variables

DESCRIPTIVE ABSTRACT:
Specifications are given for 428 new vehicles for the 2004 year. The variables recorded include price, measurements
relating to the size of the vehicle, and fuel efficiency.

SOURCE:
_Kiplinger's Personal Finance_, December 2003, vol. 57, no. 12, pp. 104-123, http:/www.kiplinger.com (permission to post on
the JSE Web site kindly granted by PARS International Corporation, 102 West 38th Street, New York, NY 10018)

VARIABLE DESCRIPTIONS:

Columns Variables
  1- 45 Vehicle Name
 47             Sports Car? (1=yes, 0=no)
 49             Sport Utility Vehicle? (1=yes, 0=no)
 51             Wagon? (1=yes, 0=no)
 53             Minivan? (1=yes, 0=no)
 55             Pickup? (1=yes, 0=no)
 57             All-Wheel Drive? (1=yes, 0=no)
 59             Rear-Wheel Drive? (1=yes, 0=no)
 61- 66 Suggested Retail Price, what the manufacturer thinks the
                vehicle is worth, including adequate profit for the
                automaker and the dealer (U.S. Dollars)
 68- 73 Dealer Cost (or "invoice price"), what the dealership pays
                the manufacturer (U.S. Dollars)
 75- 77 Engine Size (liters)
 79- 80 Number of Cylinders (=-1 if rotary engine)
 82- 84 Horsepower
 86- 87 City Miles Per Gallon
 89- 90 Highway Miles Per Gallon
 92- 95 Weight (Pounds)
 97- 99 Wheel Base (inches)
101-103 Length (inches)
105-106 Width (inches)

Values are aligned and delimited with blanks.
Missing values are denoted with *.
```

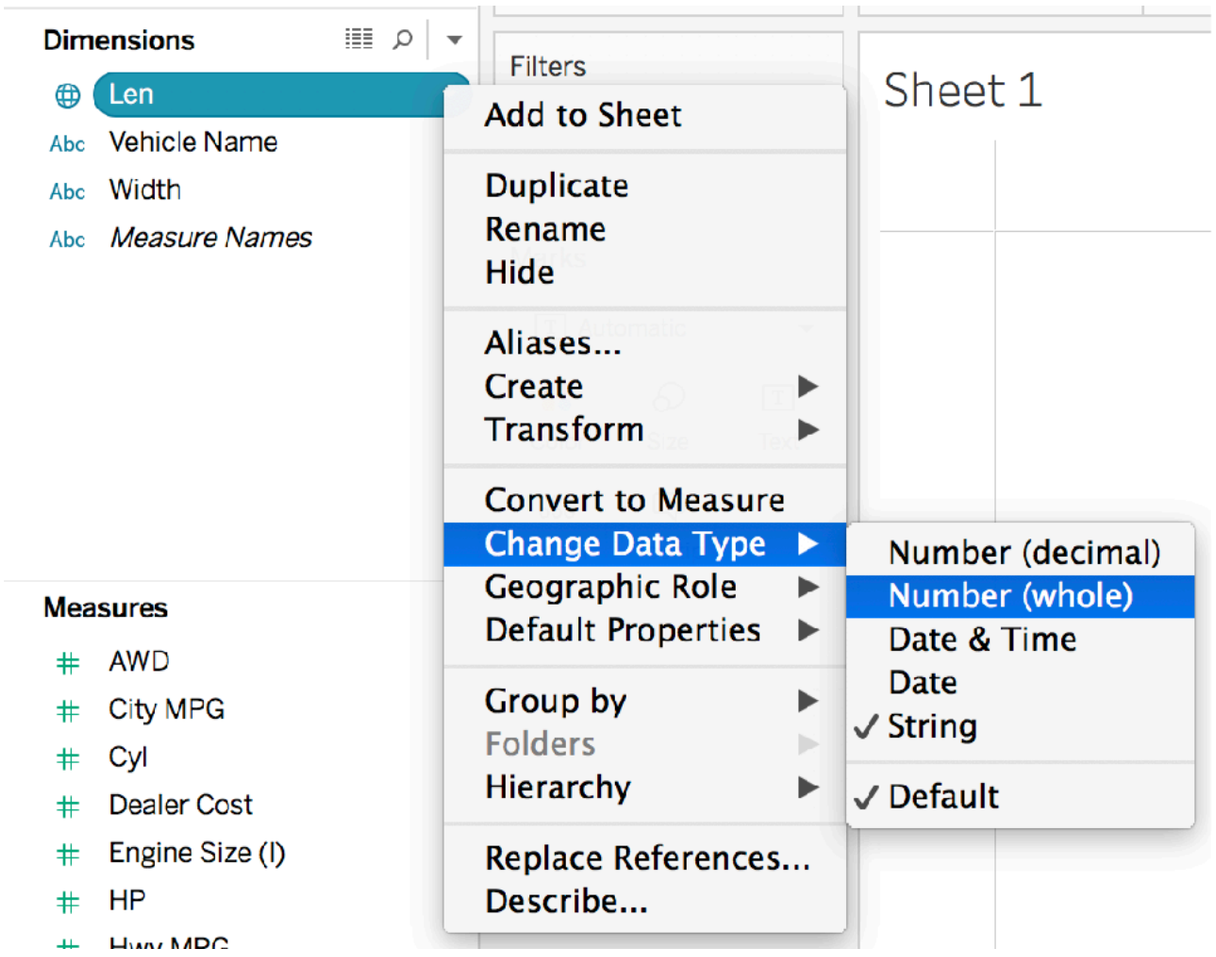
+ Extended variable names and their meaning
+ Used units
+ Special values
+ How to denote missing values

Hands On: Cars Dataset

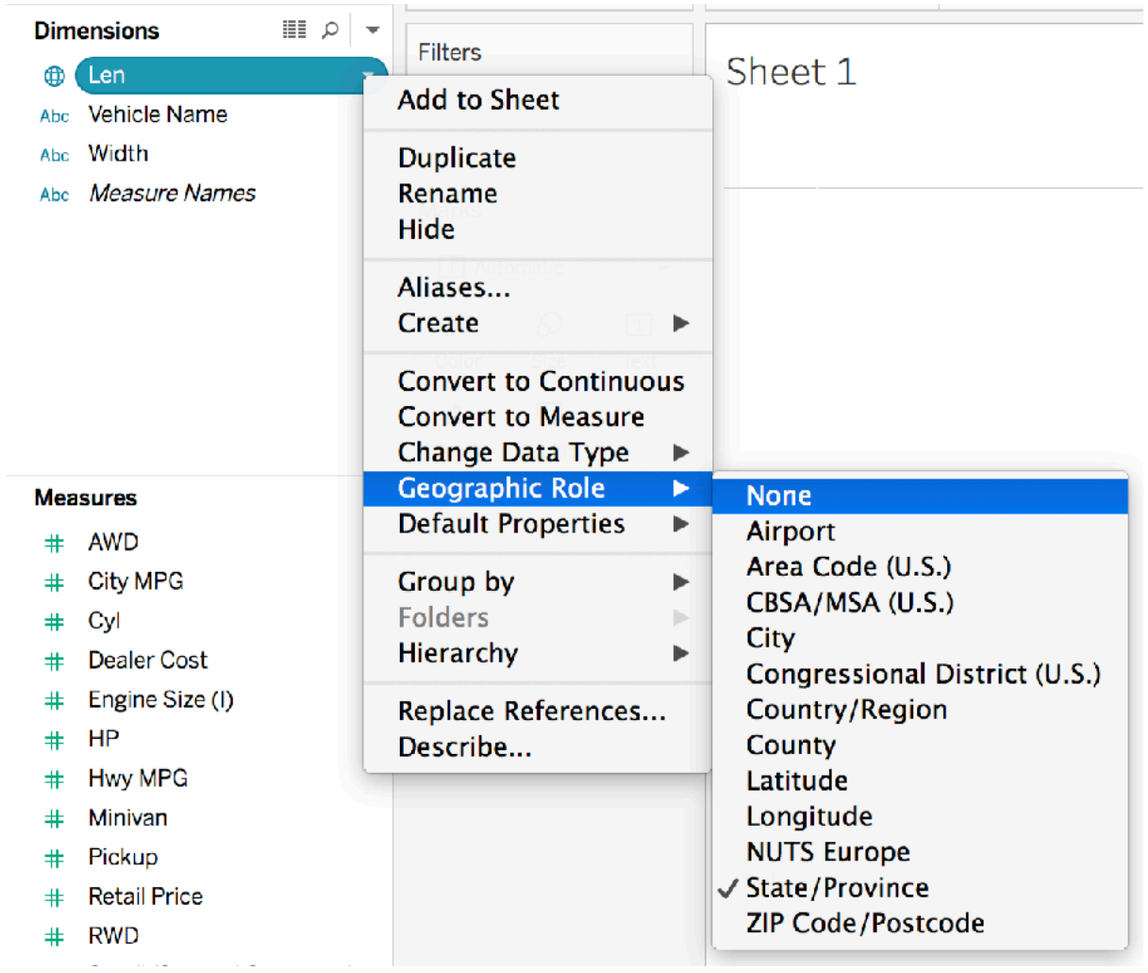
- **Steps:**
  - **Data organization:**
    - if the data type is appropriate - **Fix data types**!
    - **Dimension or Measure?** For each variable decide wether it should be considered a dimension or a measure.
    - Some variables that could facilitate our analysis are missing and are dispersed or included in other variables - **Create additional variables**!
    - Some Dimensions are appropriate to **create hierarchies**. For instance (**Brand**,**Model**), just like **Year**/**Quarter**/**Month** for Date dimensions.

Data Organisation

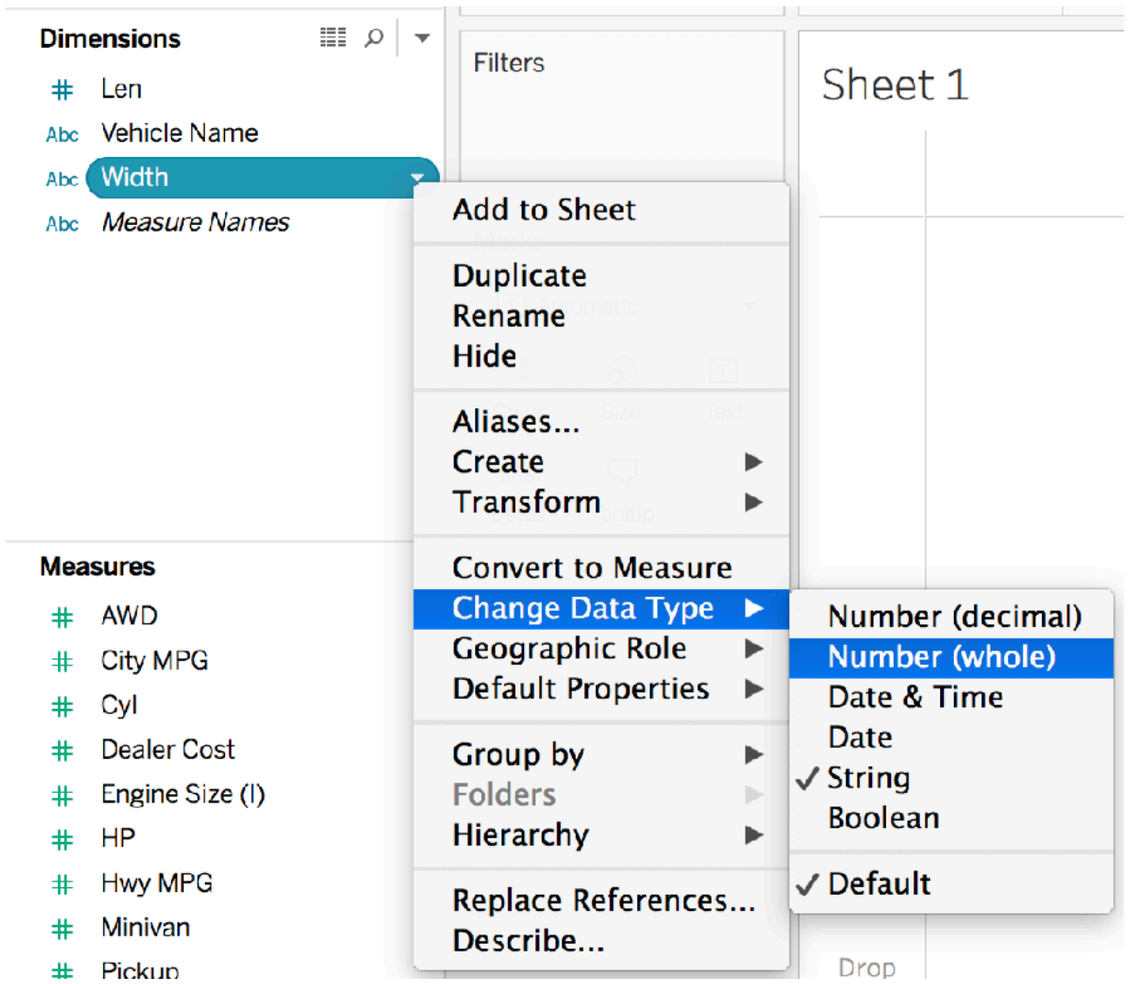
Fix Data Types



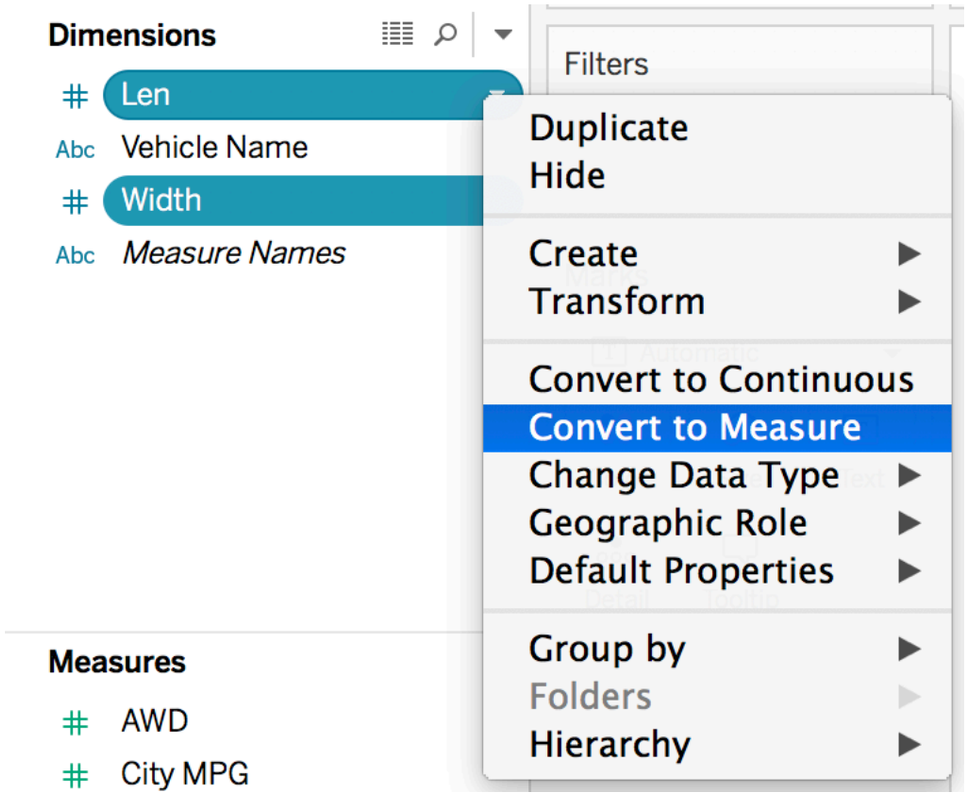
Fix Data Types



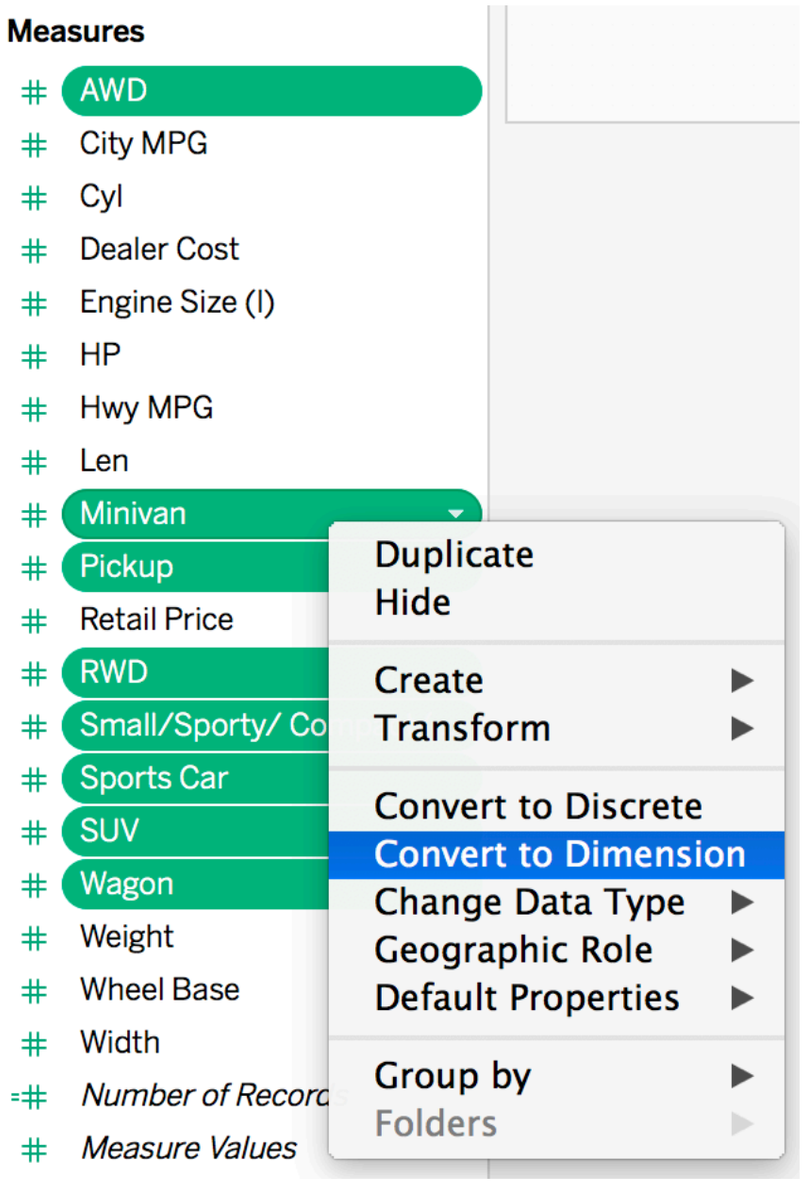
Fix Data Types



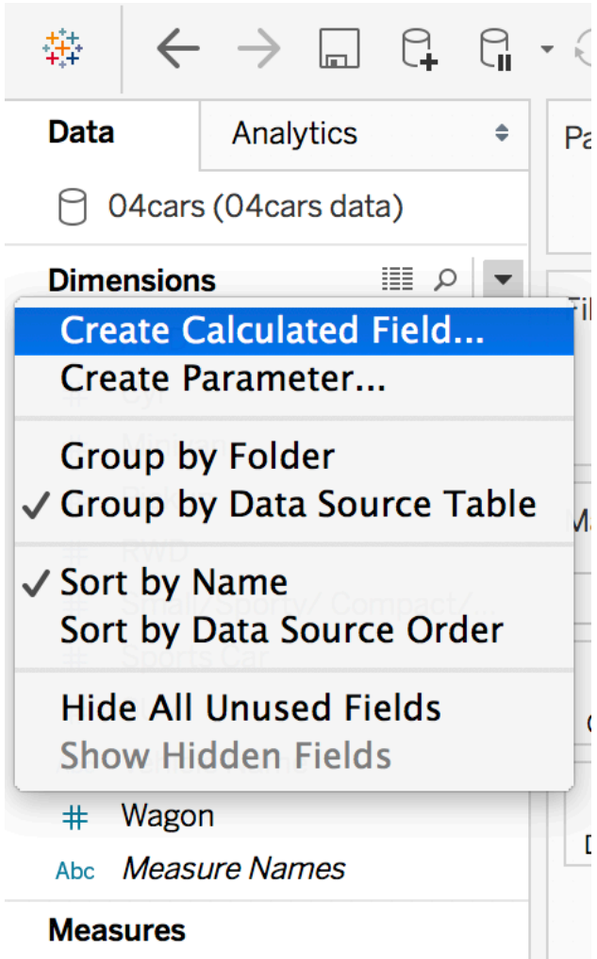
Dimension or Measure?



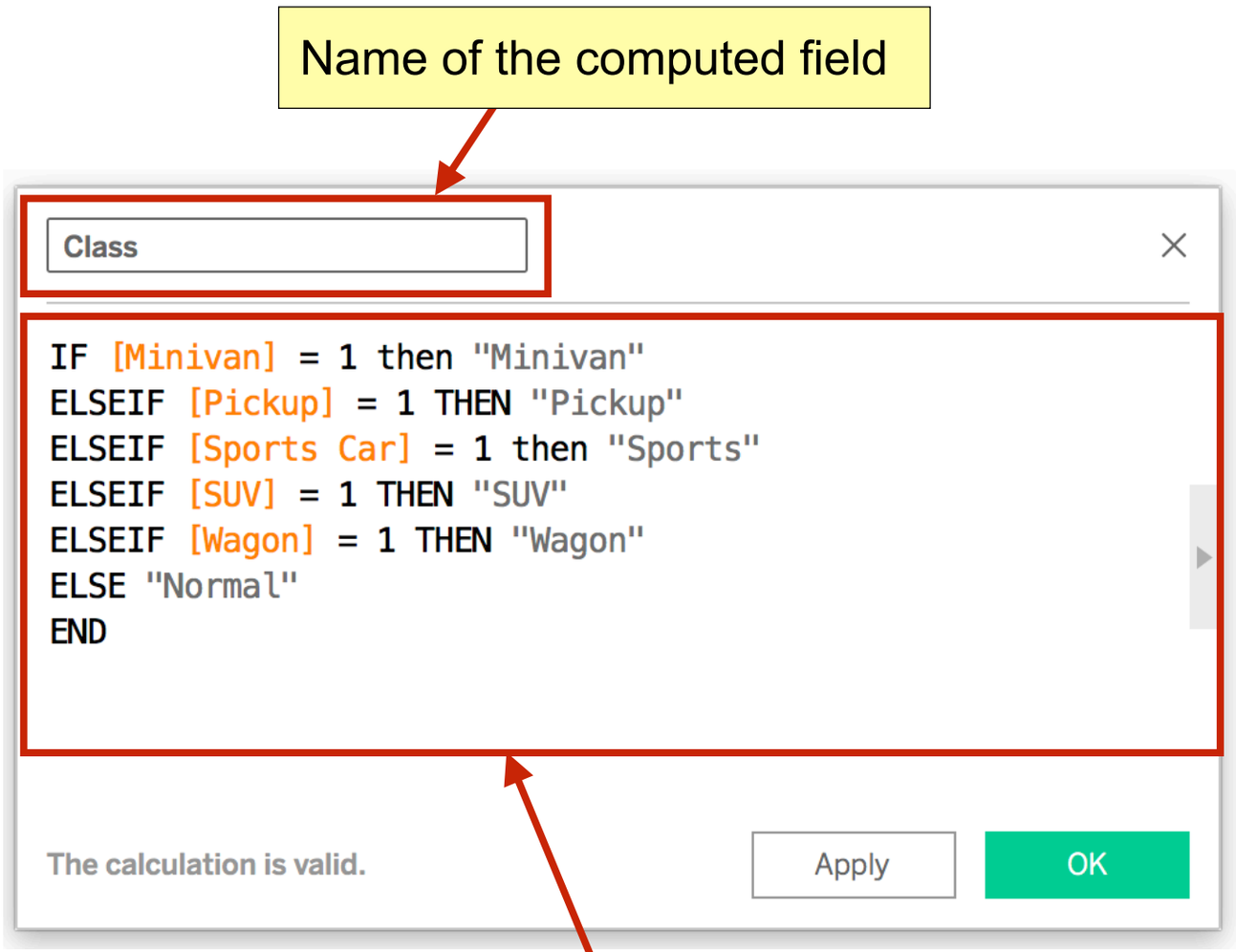
Dimension or Measure?



Create Additional Variables

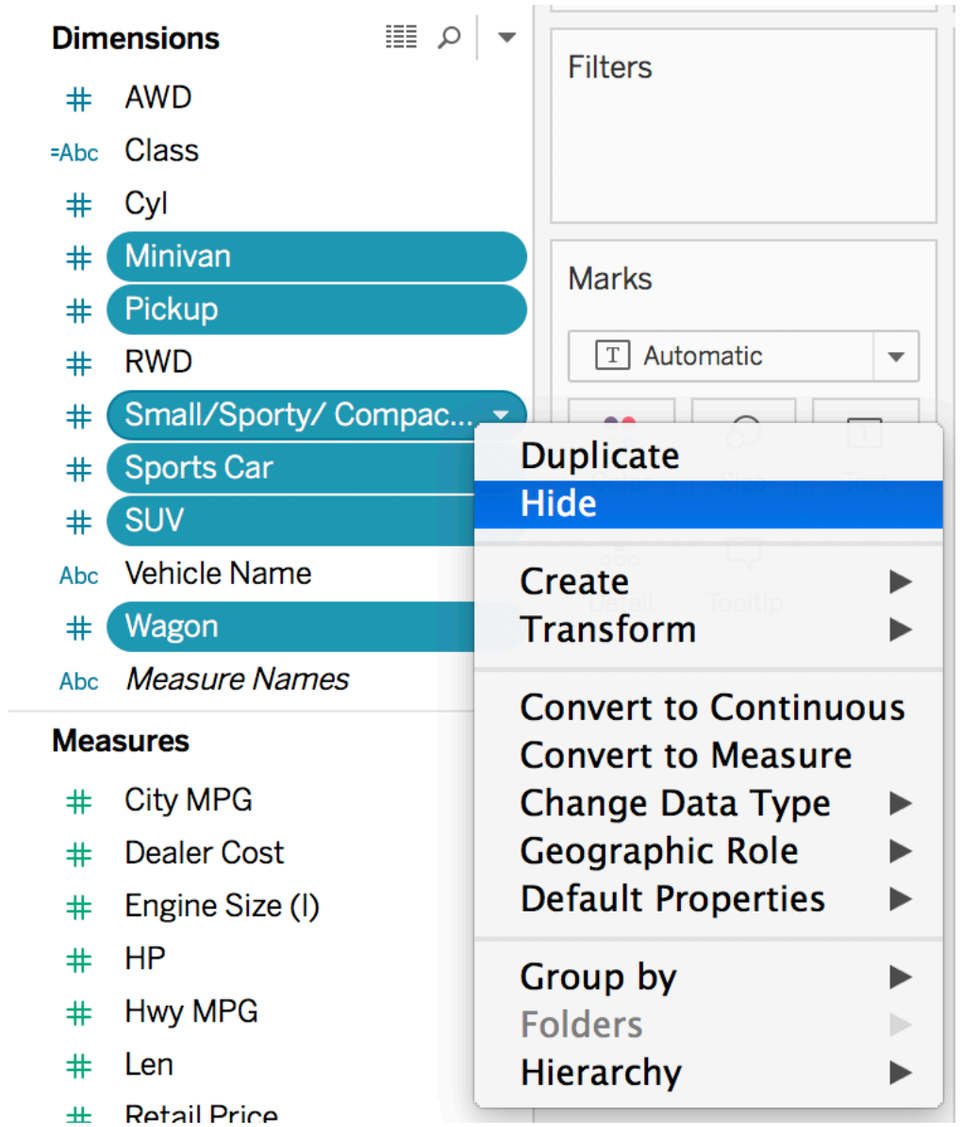


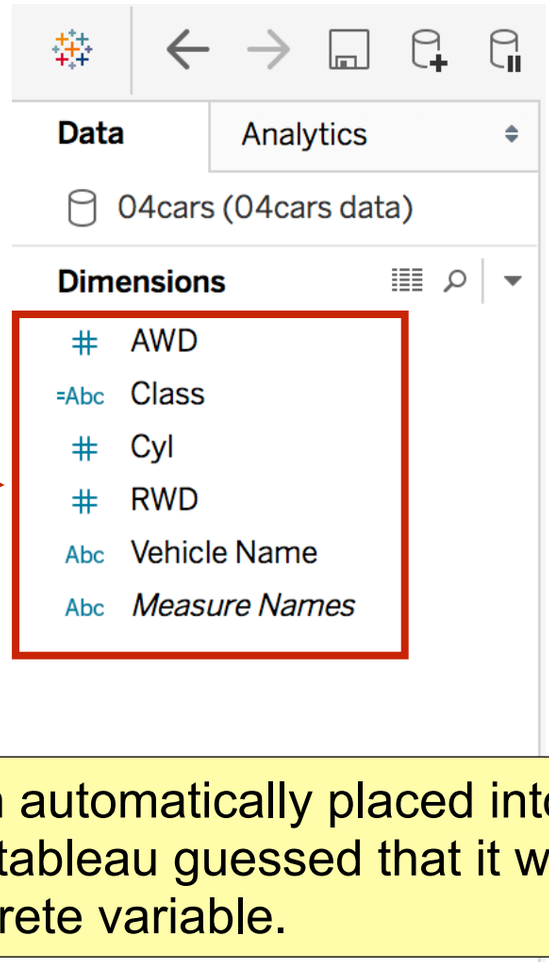
Name of the computed field



Code that computes the field's value

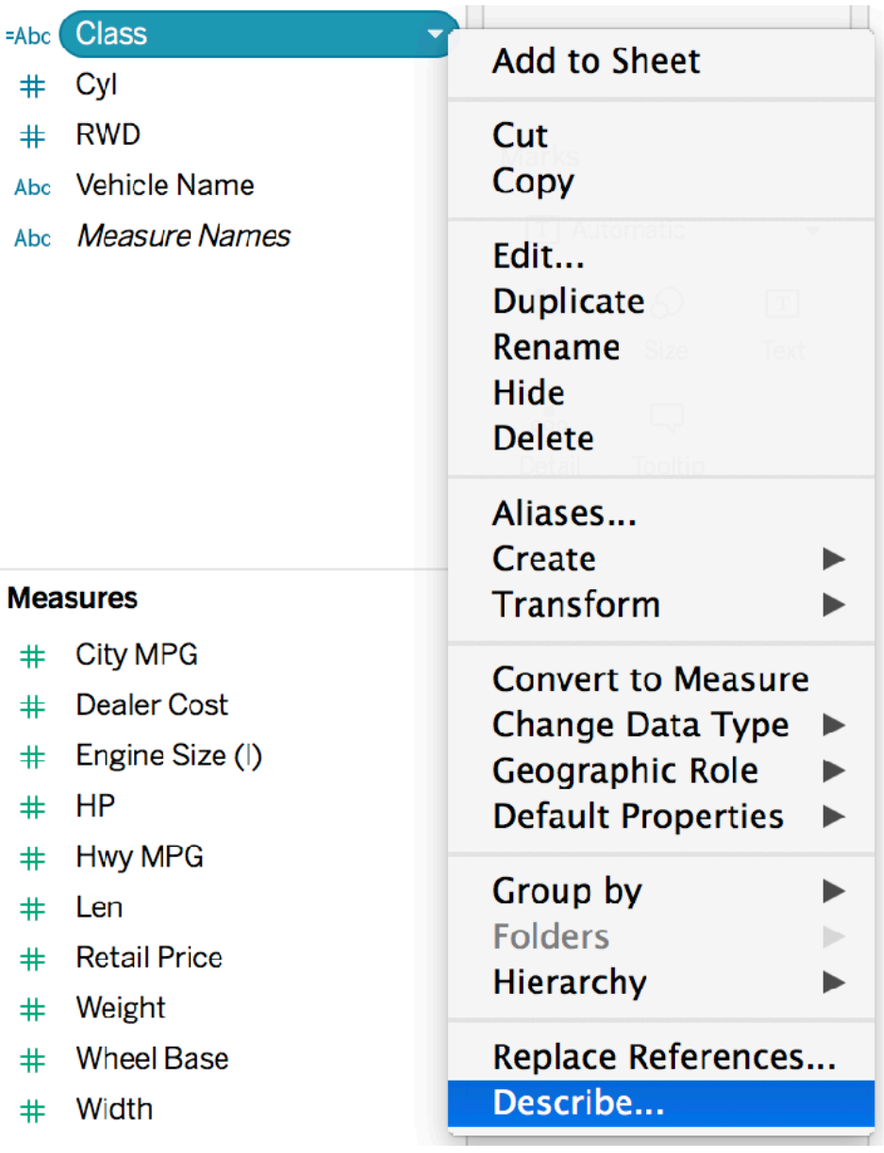
Create Additional Variables

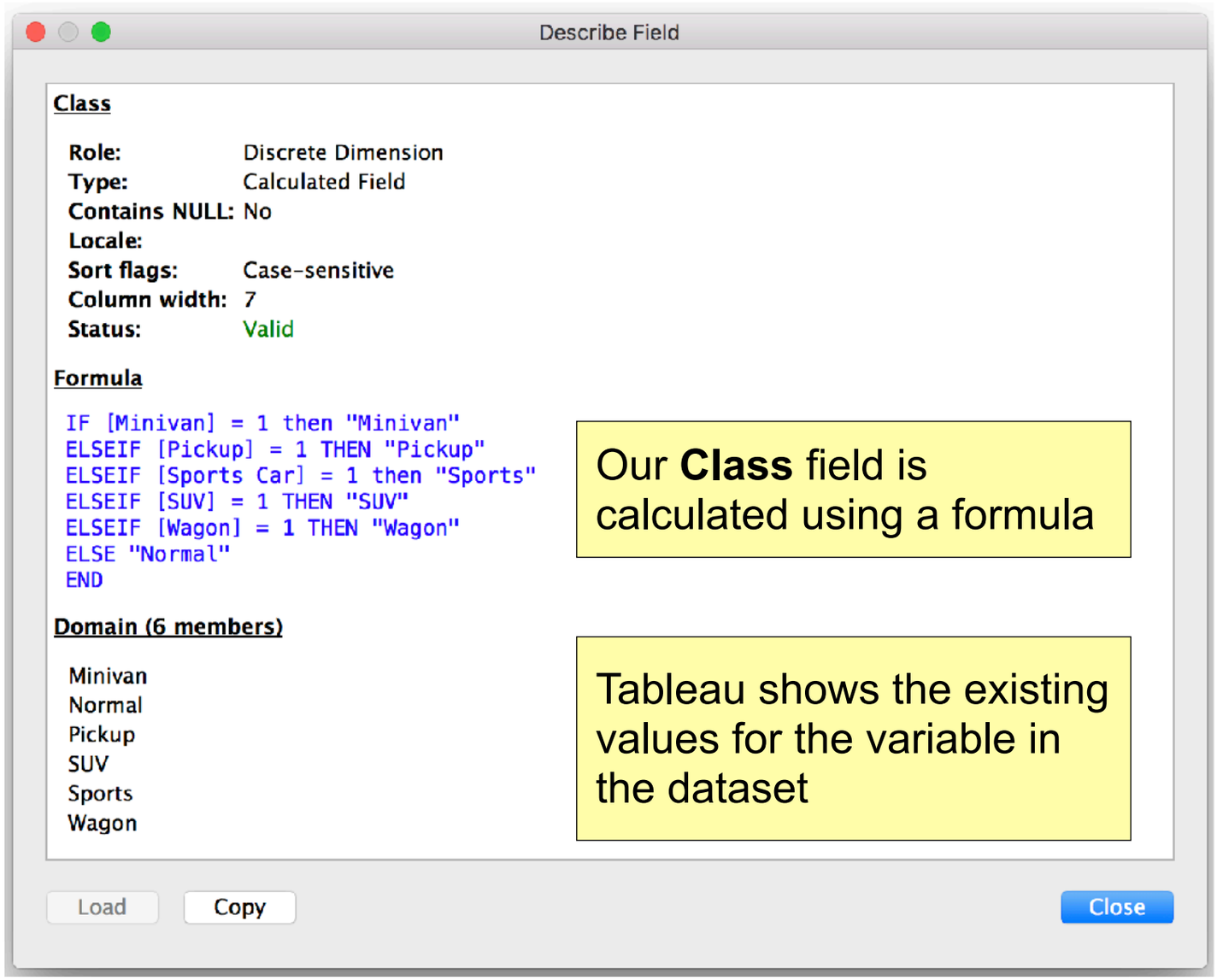




Class has been automatically placed into Dimensions as tableau guessed that it was a categorical/discrete variable.

Create Additional Variables



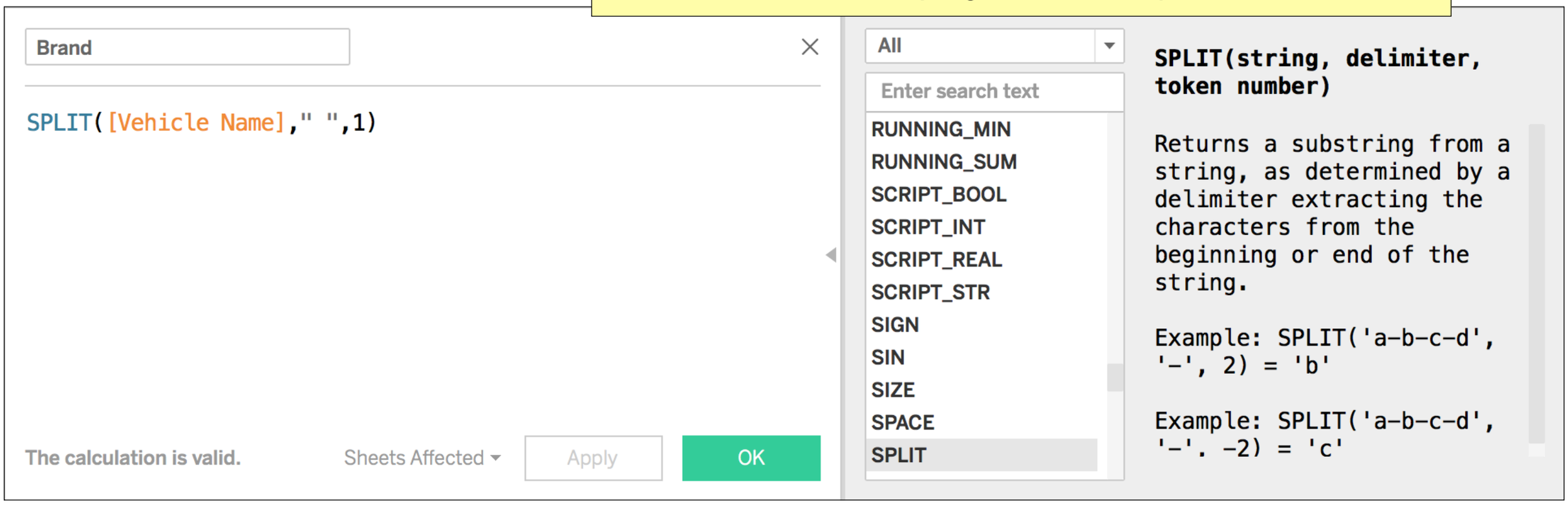


Our **Class** field is calculated using a formula

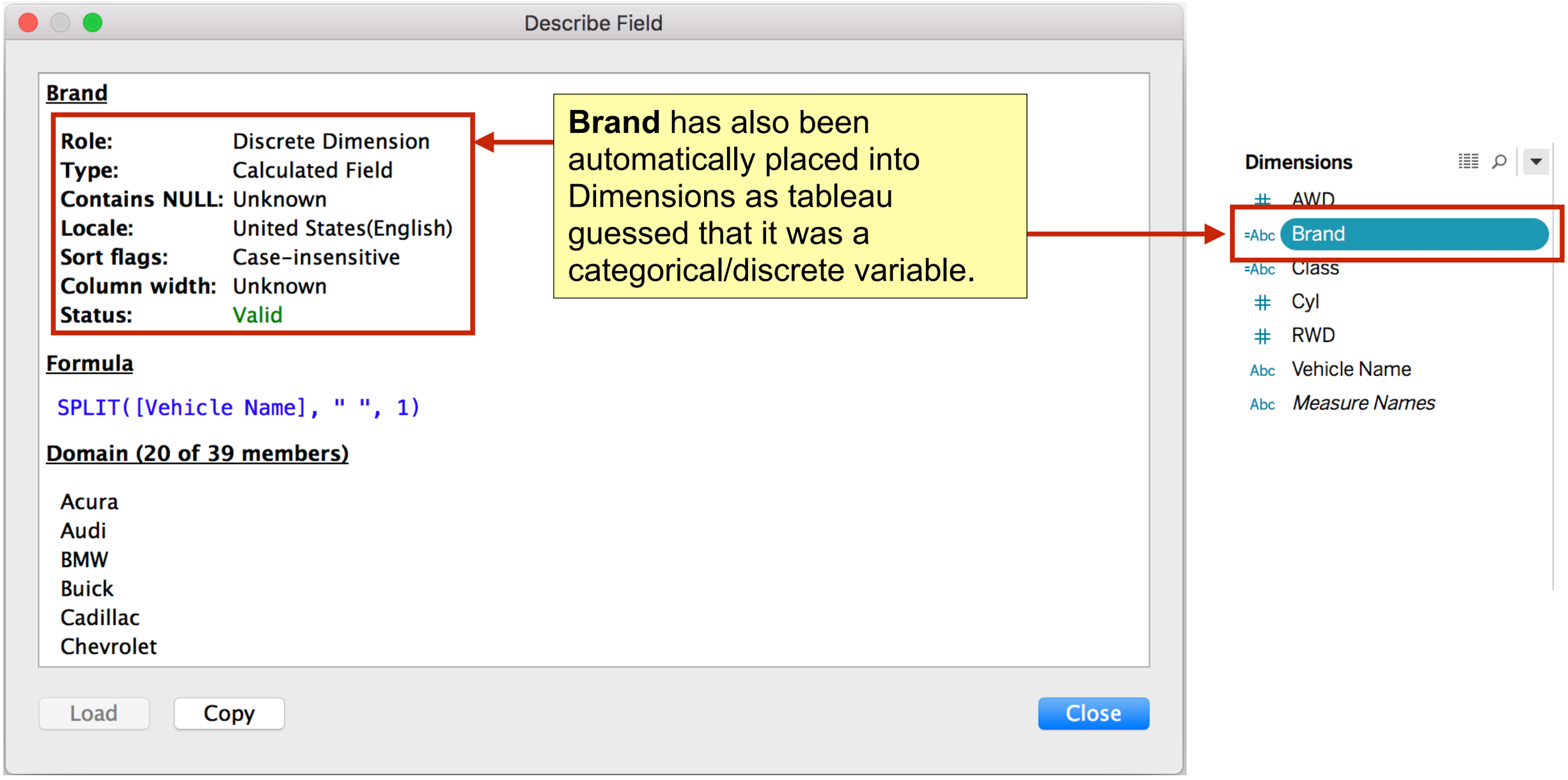
Tableau shows the existing values for the variable in the dataset

Create Additional Variables

Brand is also a computed field that results from **splitting** the “Vehicle Name” using the **space character** as a separator and keeping the first part. For “Model” use the same formula but keeping the second part instead.

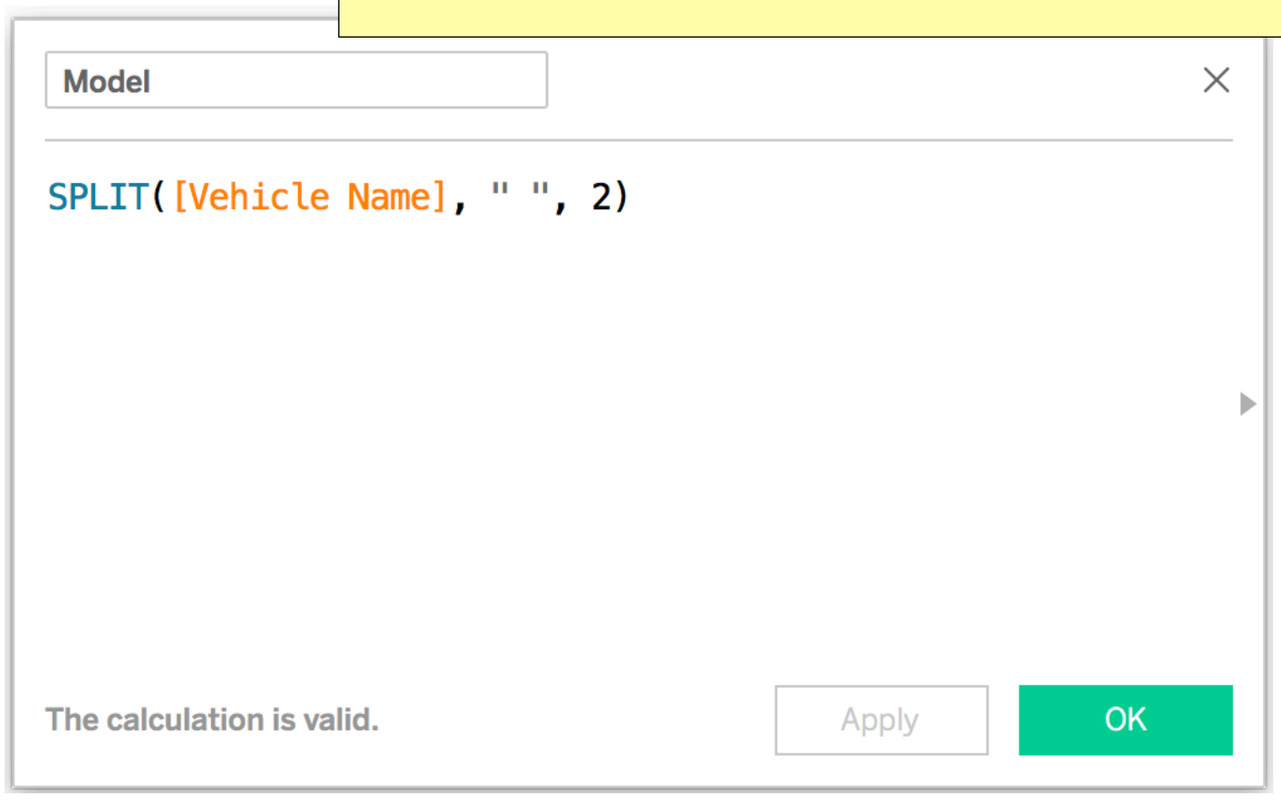


Create Additional Variables



Create Additional Variables

For “Model” use the same formula as for the “Brand” but keeping the second part instead.



Vehicle Name

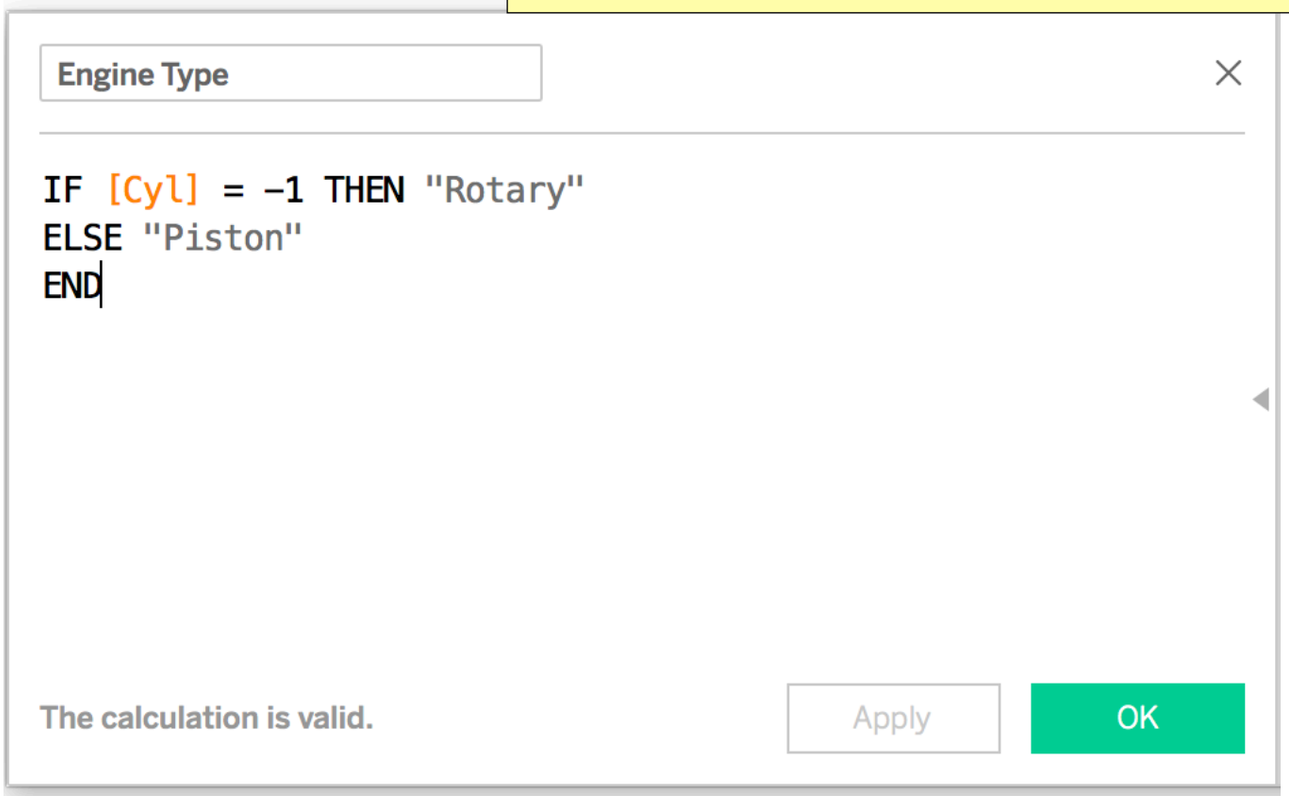
Role:	Discrete Dimension
Type:	Database column
Remote column:	[04cars].[Vehicle Name]
Remote type:	Unicode character string
Contains NULL:	Unknown
Locale:	United States(English)
Sort flags:	Case-insensitive
Column width:	Unknown
Status:	Valid

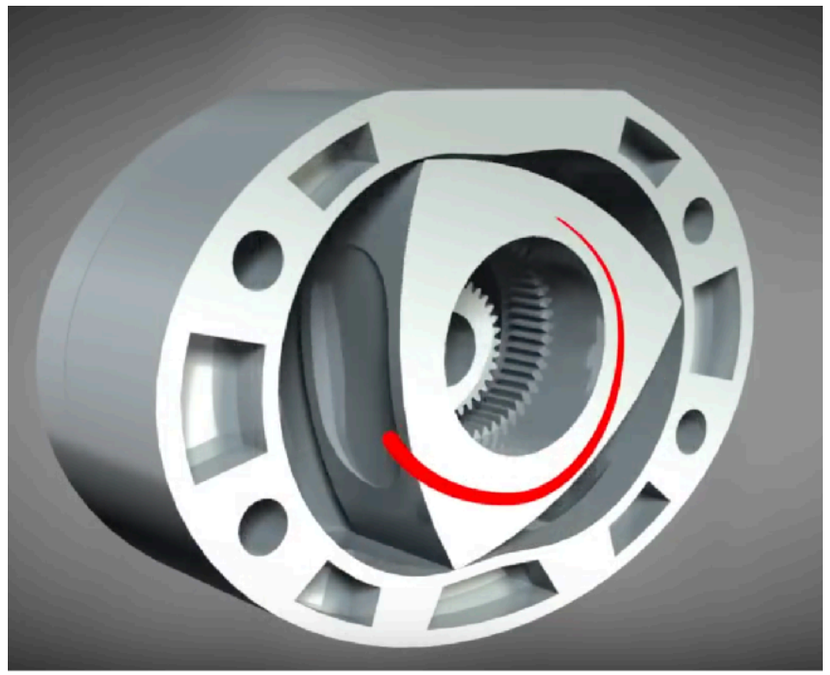
Domain (20 of 425 members)

Acura 3.5 RL 4dr
Acura 3.5 RL w/Navigation 4dr
Acura MDX
Acura NSX coupe 2dr manual S
Acura RSX Type S 2dr
Acura TL 4dr
Acura TSX 4dr

Create Additional Variables

A rotary engine is a type of combustion engine that does not use pistons inside cylinder chambers. It has been used by Mazda in some of their cars until 2012 (see https://en.wikipedia.org/wiki/Mazda_Wankel_engine)





https://youtu.be/josJhz8VS8A

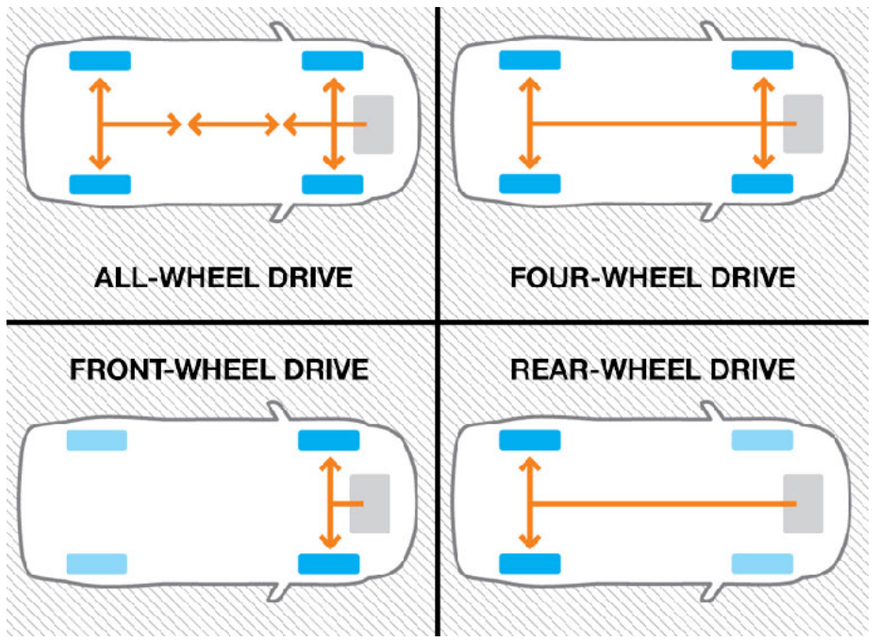
Create Additional Variables

FWD - Front wheel drive
AWD - All wheel drive
RWD - Rear wheel drive

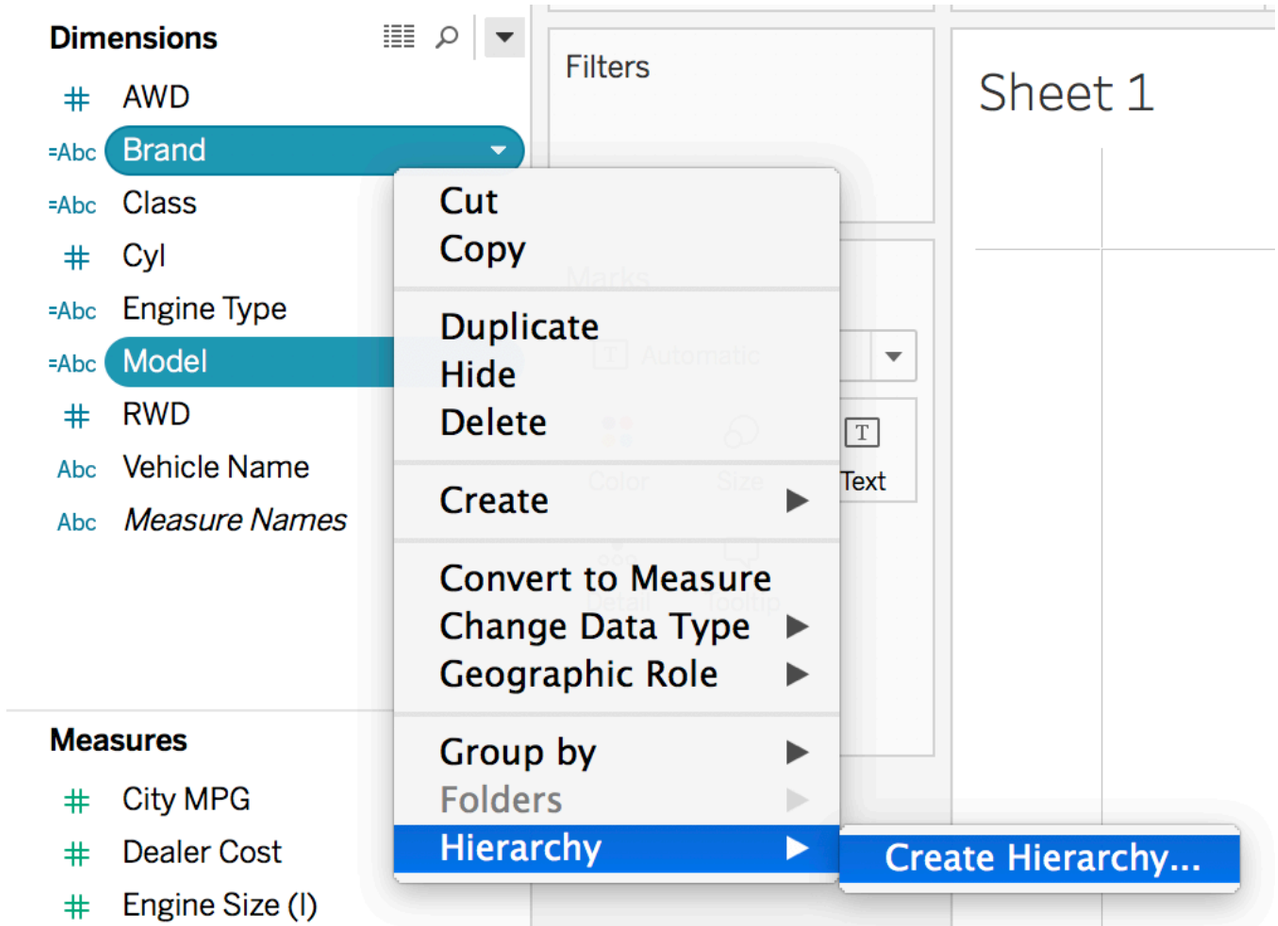
Drive Train

```
IF [AWD] = 1 THEN "AWD"
ELSEIF [RWD] = 1 THEN "RWD"
ELSE "FWD"
END
```

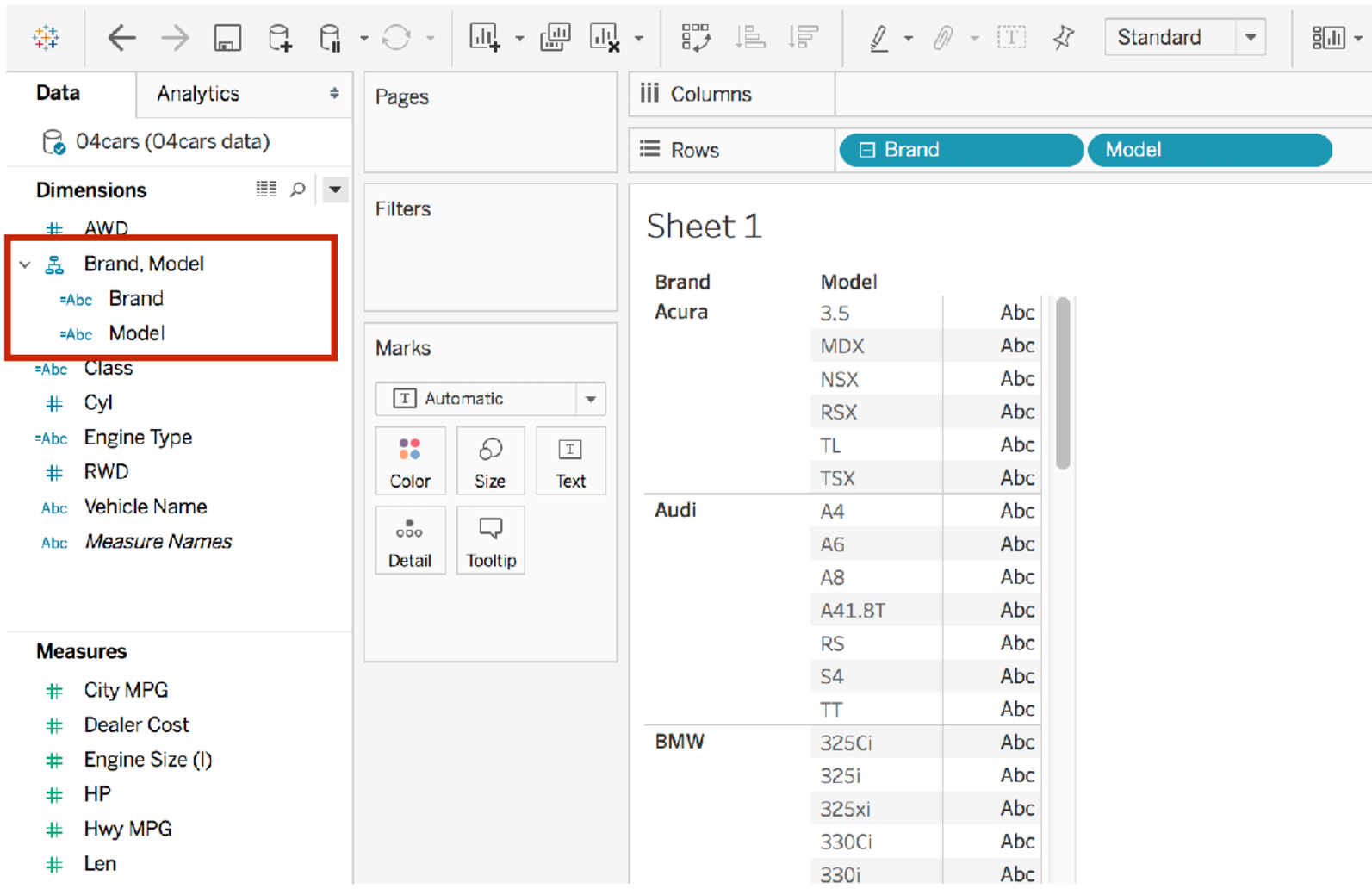
The calculation is valid. Sheets Affected Apply OK



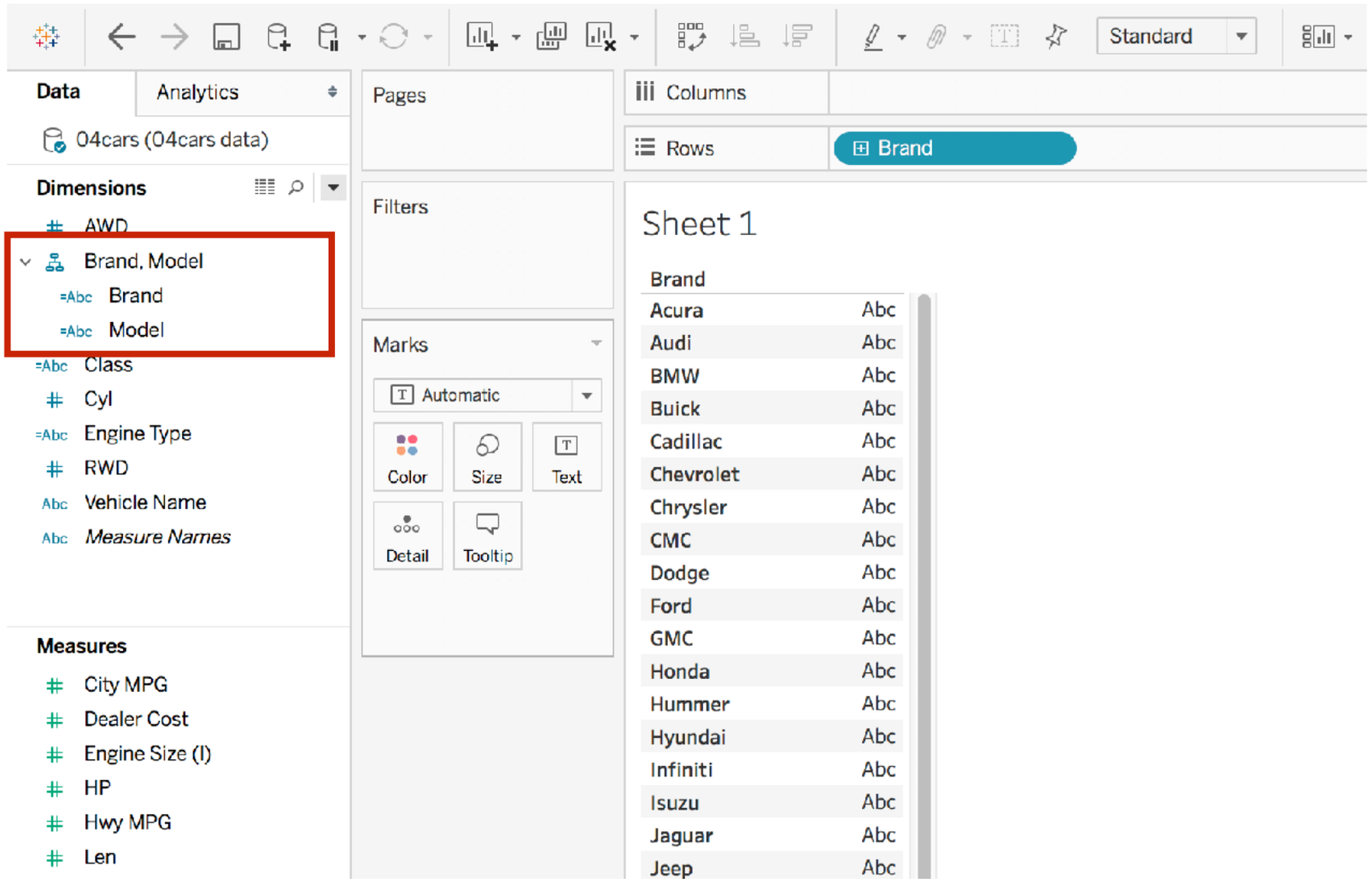
Create Hierarchies



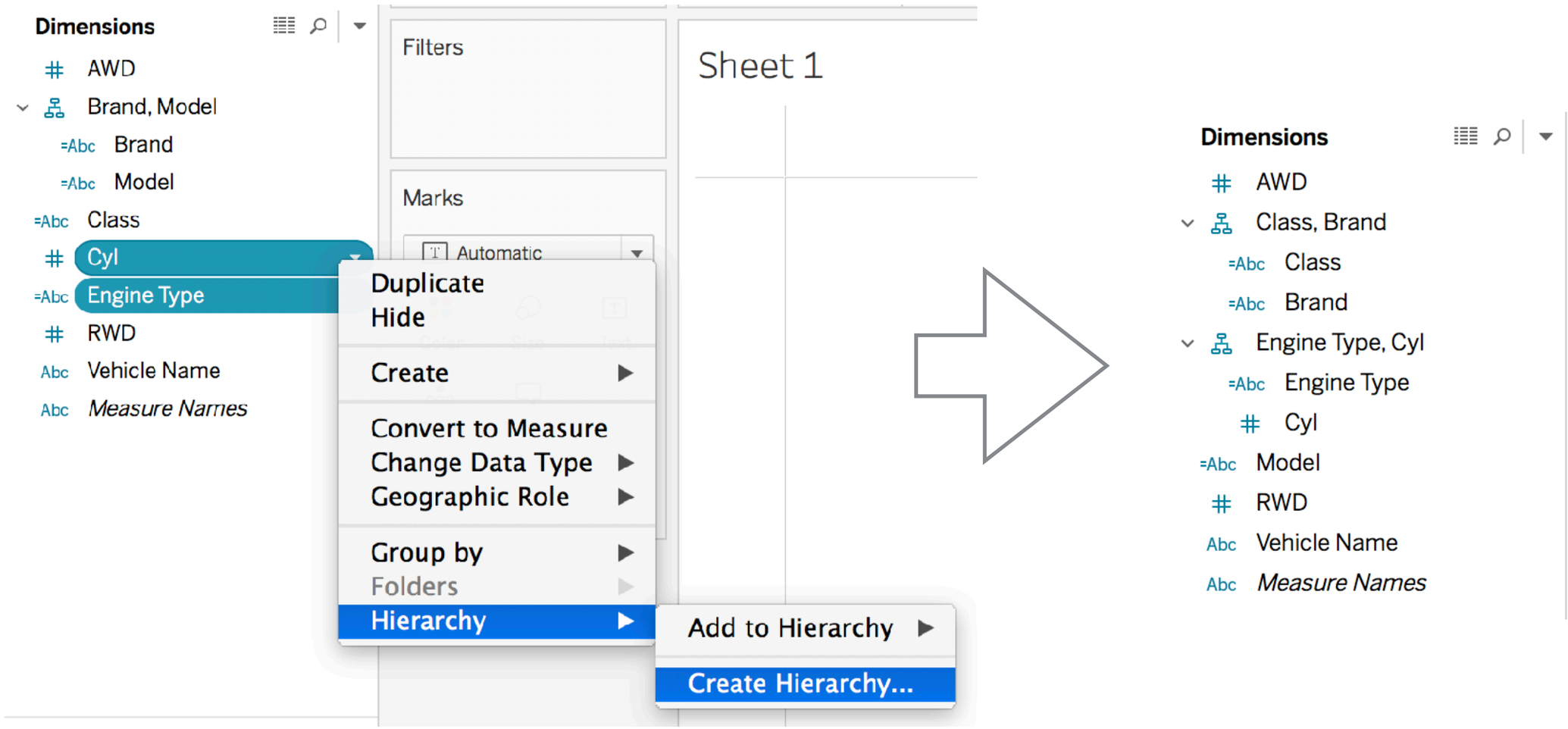
Create Hierarchies



Create Hierarchies



Create Hierarchies



Hands On: Cars Dataset

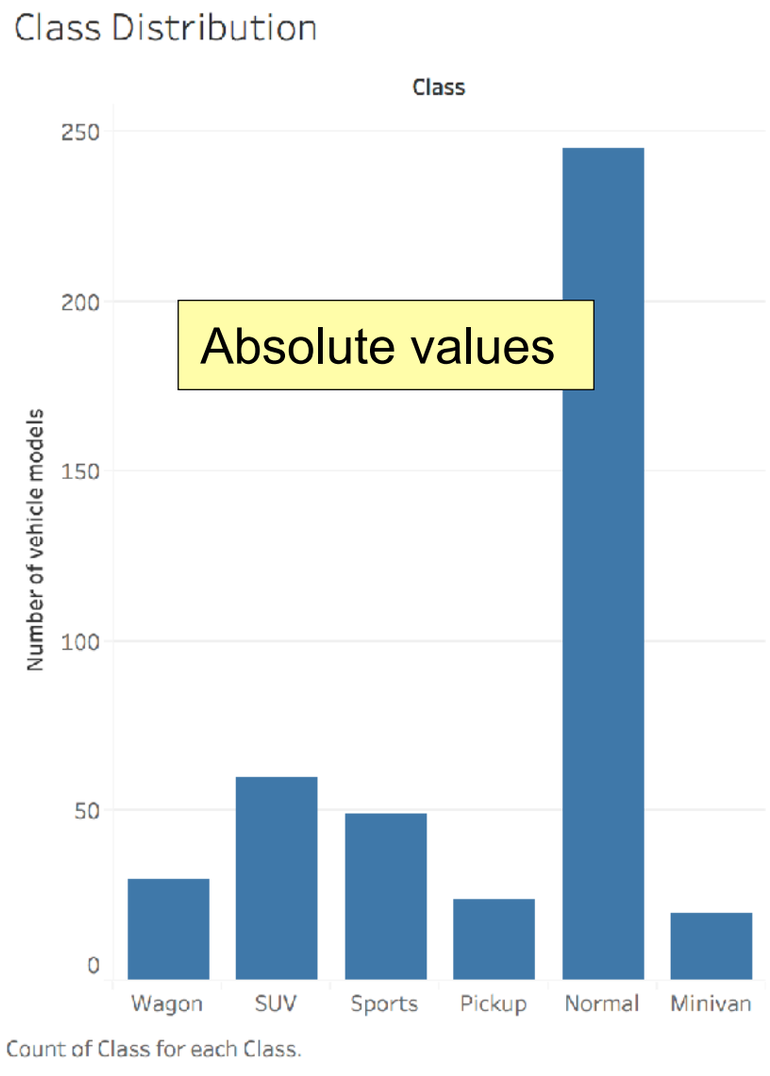
- **Steps:**
  - **Explore the Data:**
    - How are the variables distributed? Show the **distribution of** each **variable**!
    - Show the joint distribution of variables
    - Show your distribution broken down by another dimension

Explore the Data

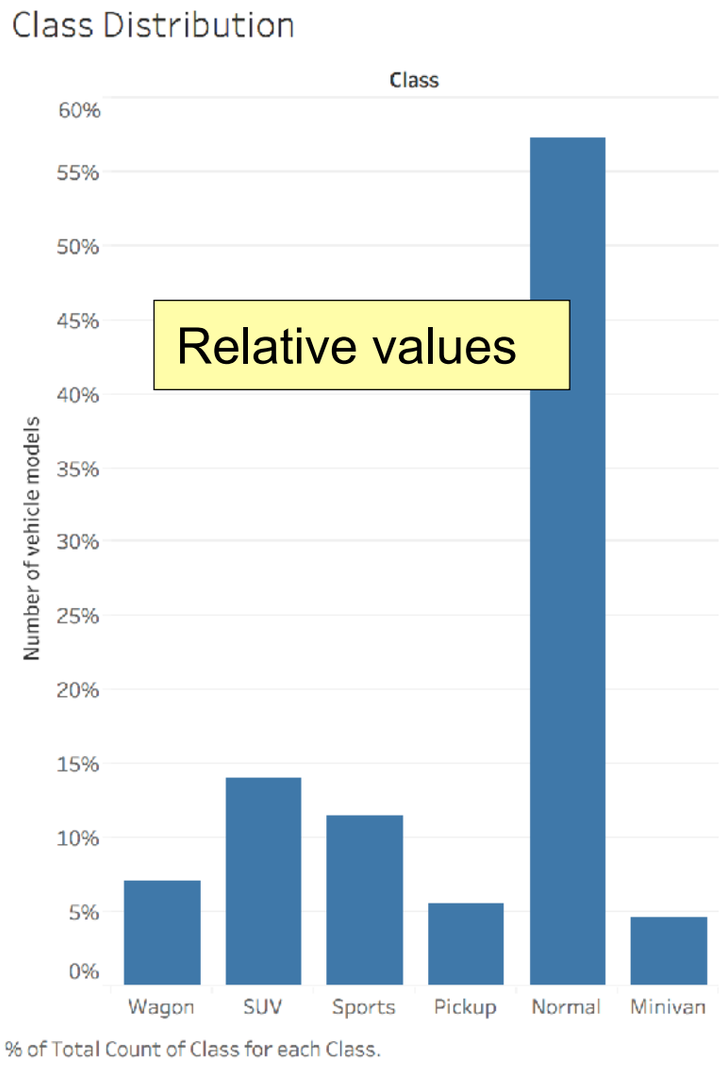
Explore the Data: Single Variable

Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Make a **bar chart** of each **discrete variable** counting the occurrences



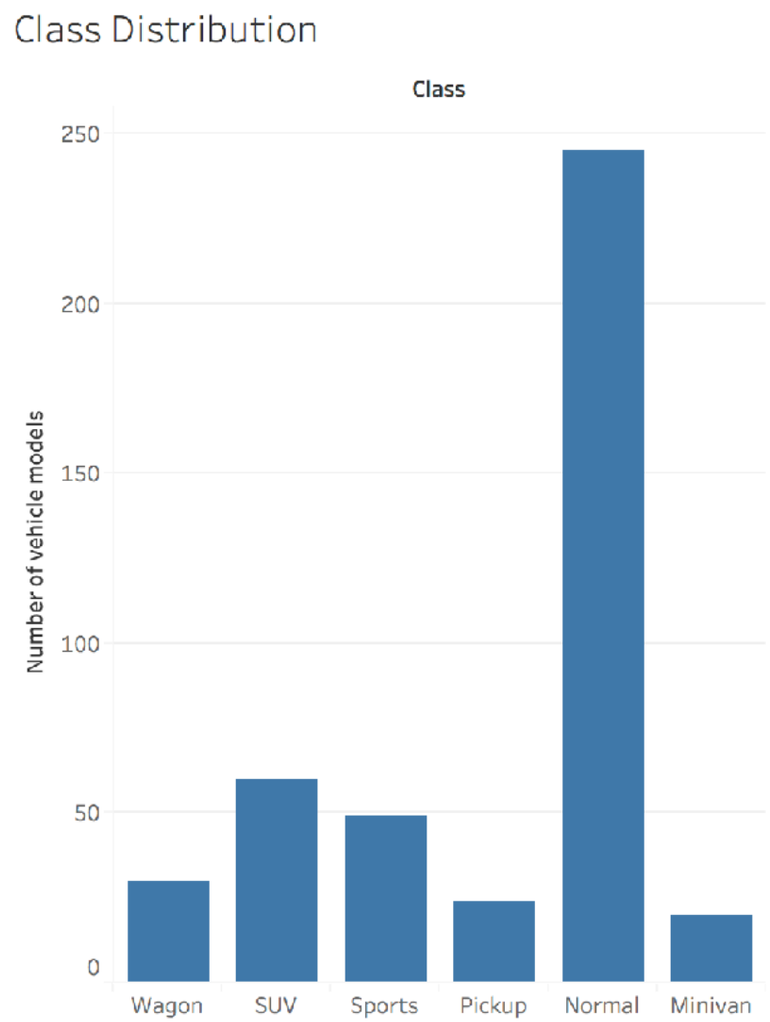
Absolute values

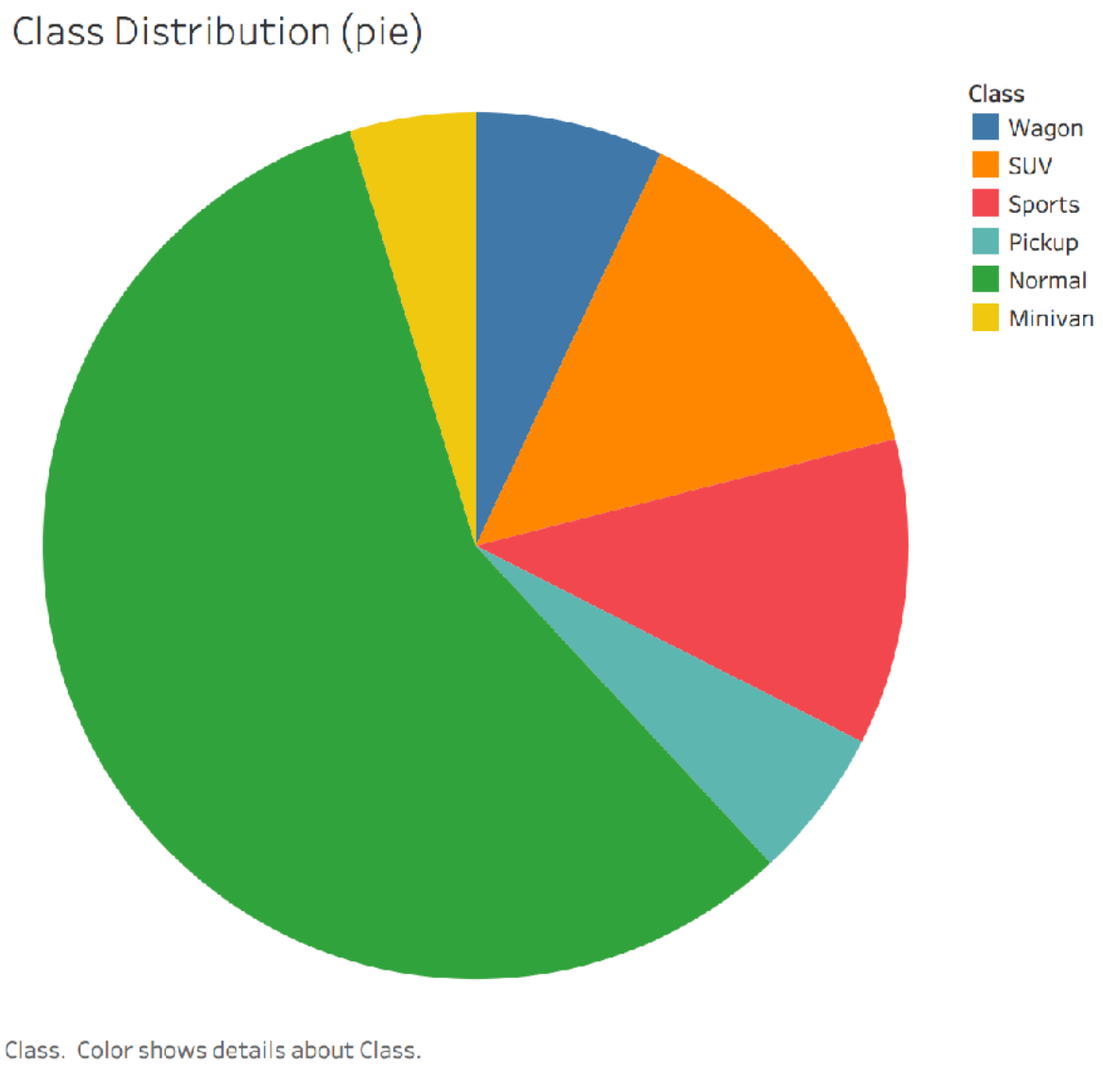


Relative values

Distribution of a (categorical) variable

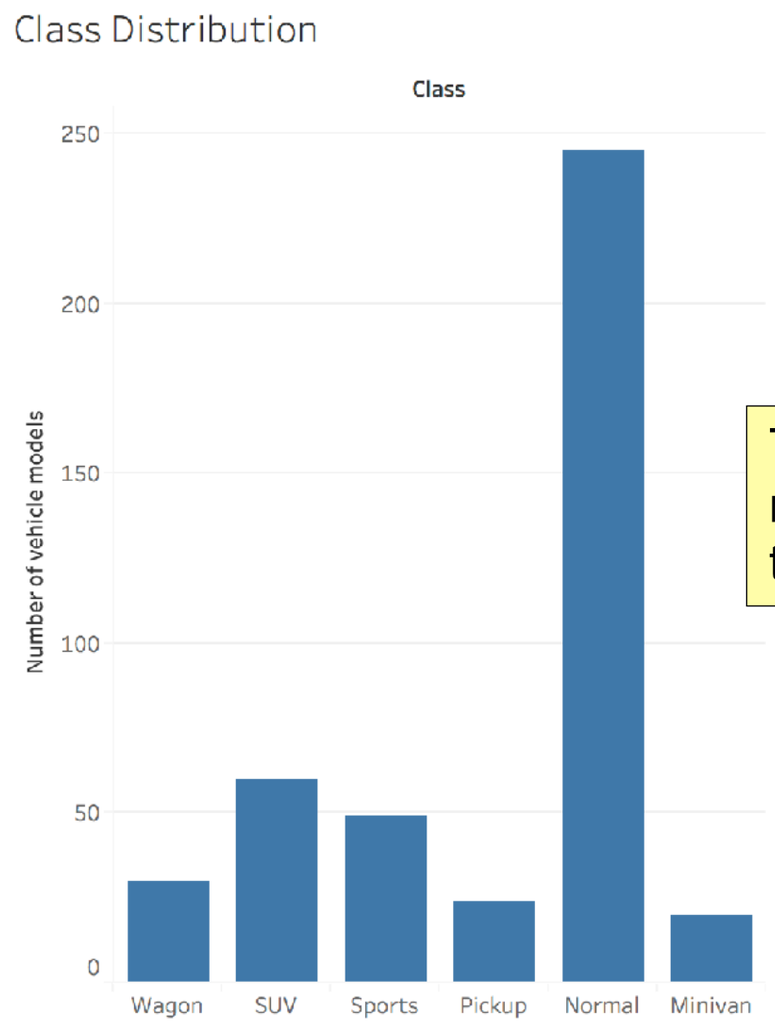
- **Compute the distribution of each variable**
  - A bar chart can be an alternative for a relative distribution

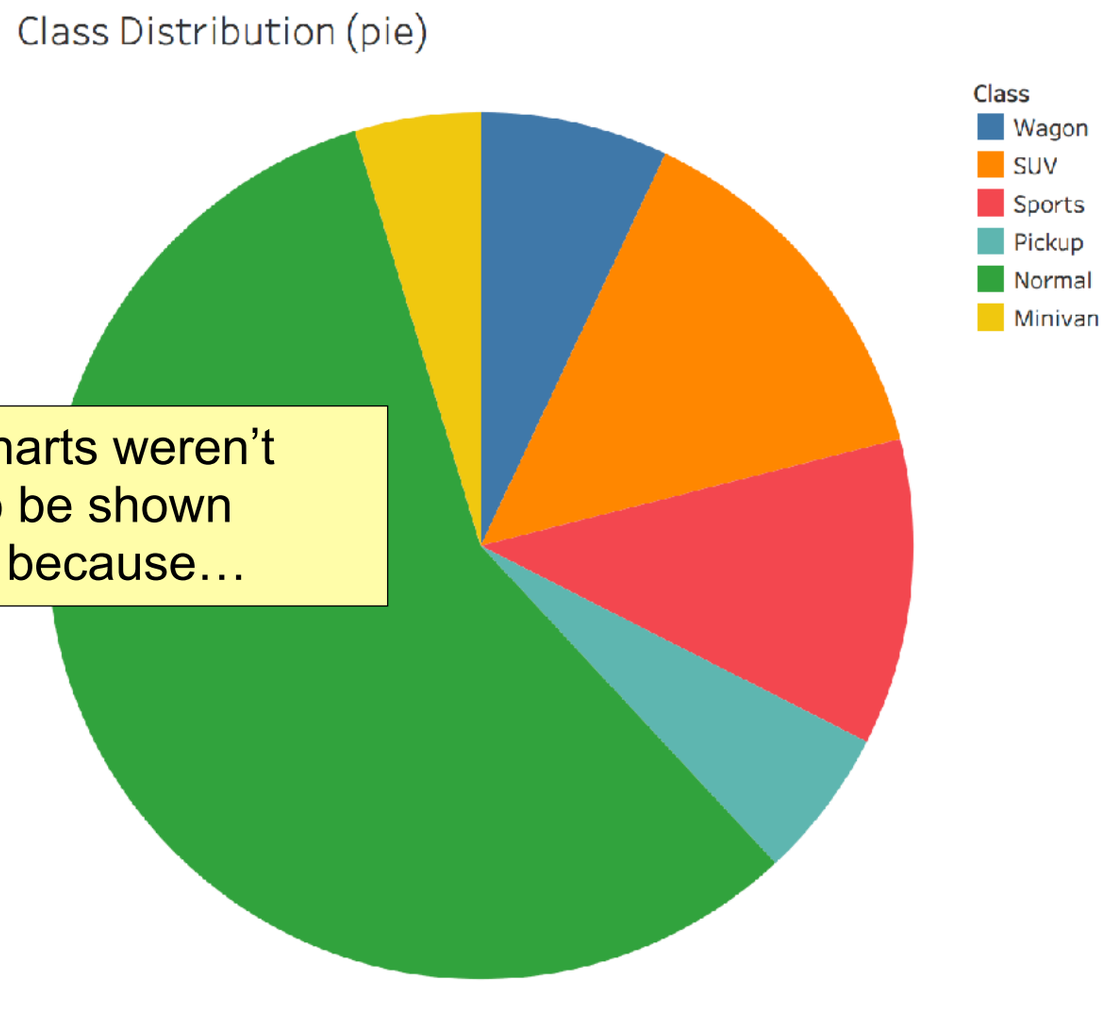




Distribution of a (categorical) variable

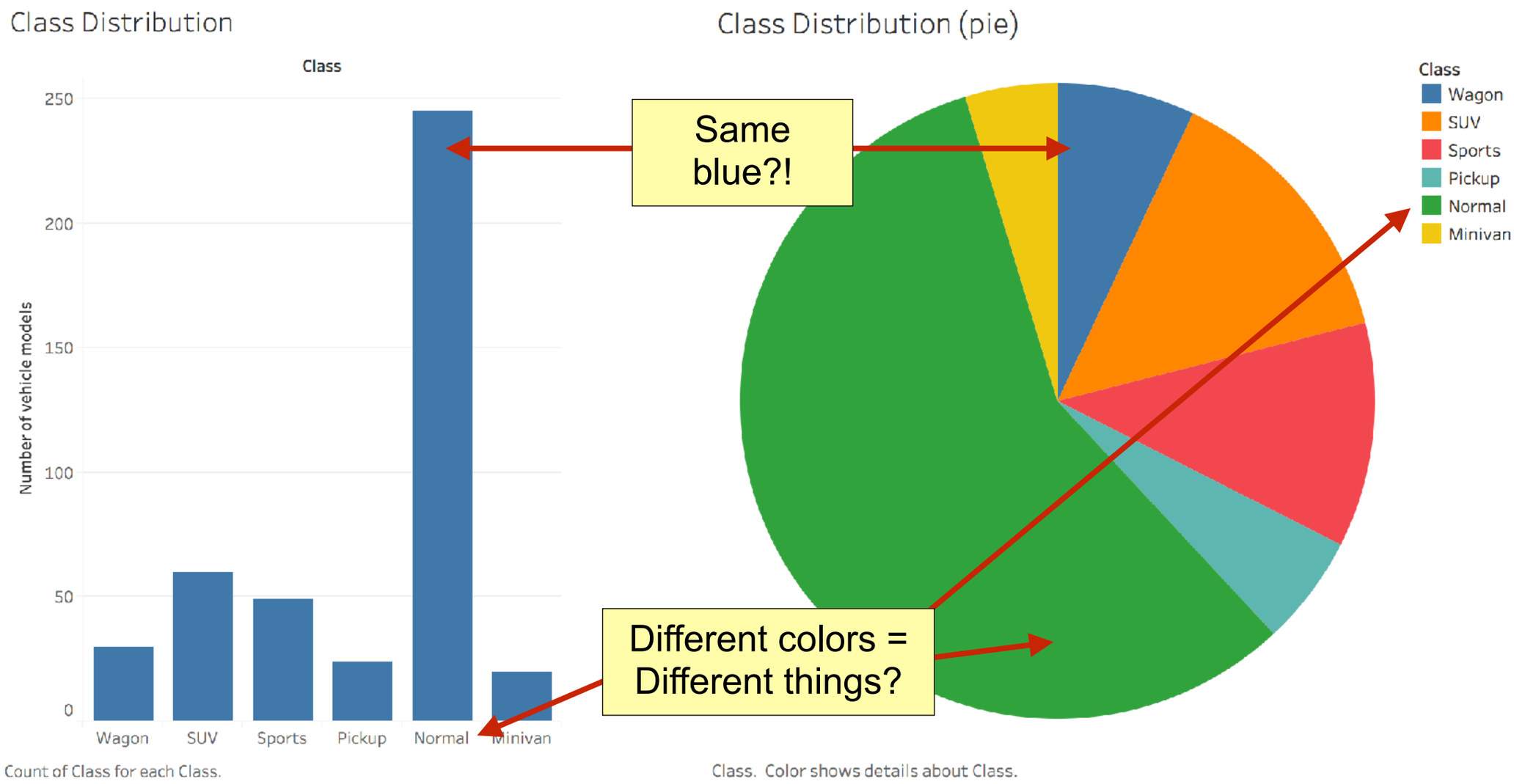
- **Compute the distribution of each variable**
  - A bar chart can be an alternative for a relative distribution





Distribution of a (categorical) variable

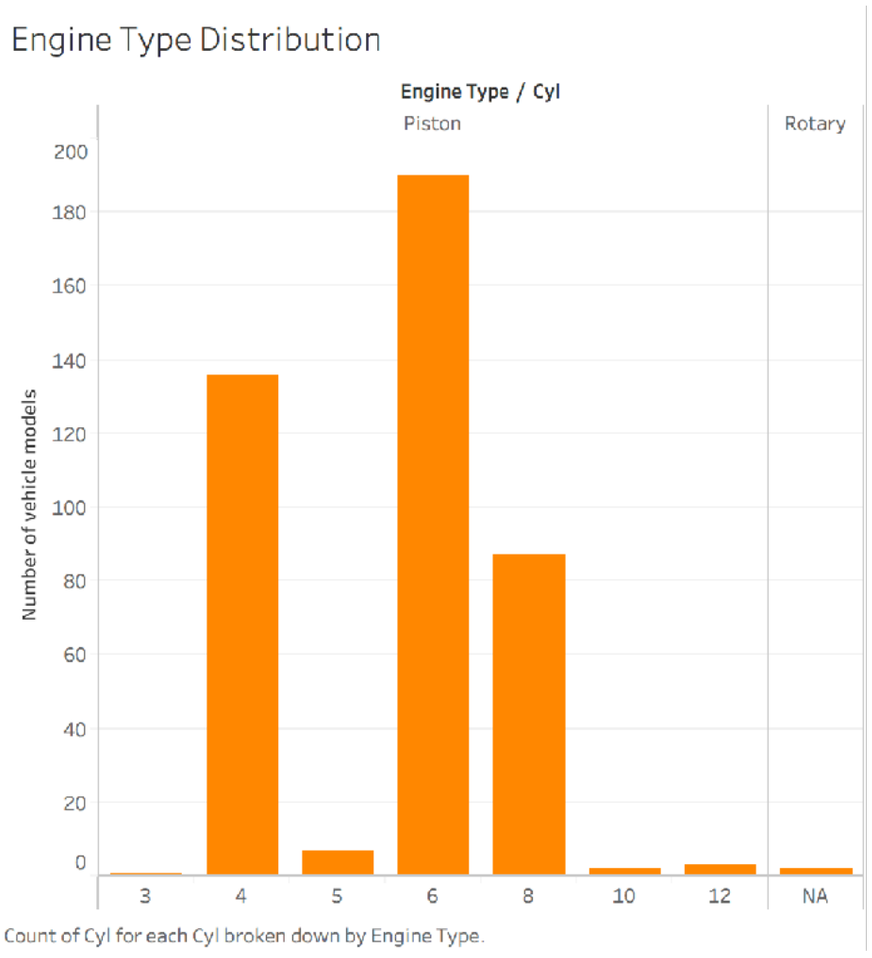
- **Compute the distribution of each variable**
  - A bar chart can be an alternative for a relative distribution

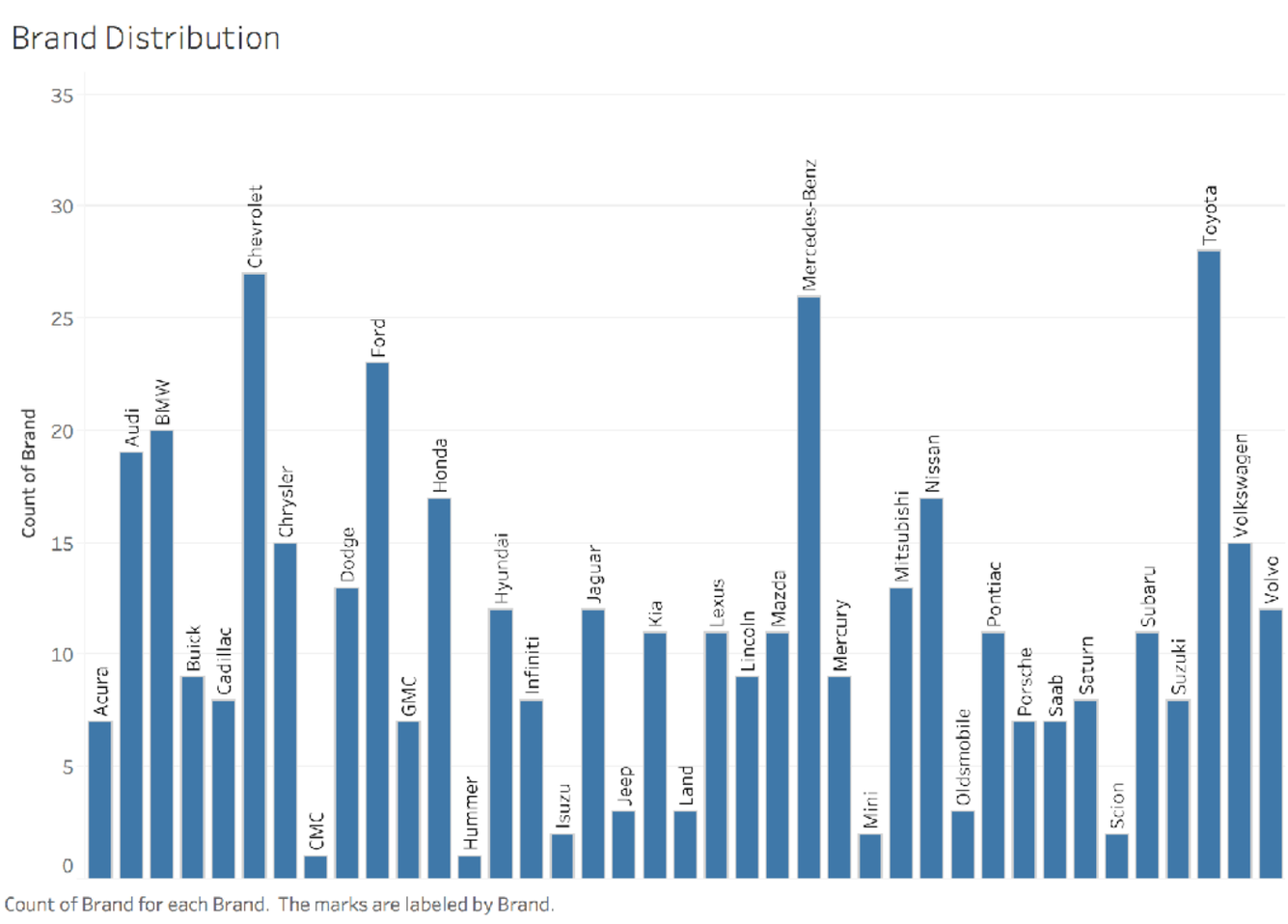


Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Make a **bar chart** of each **discrete variable** counting the occurrences

Additional examples

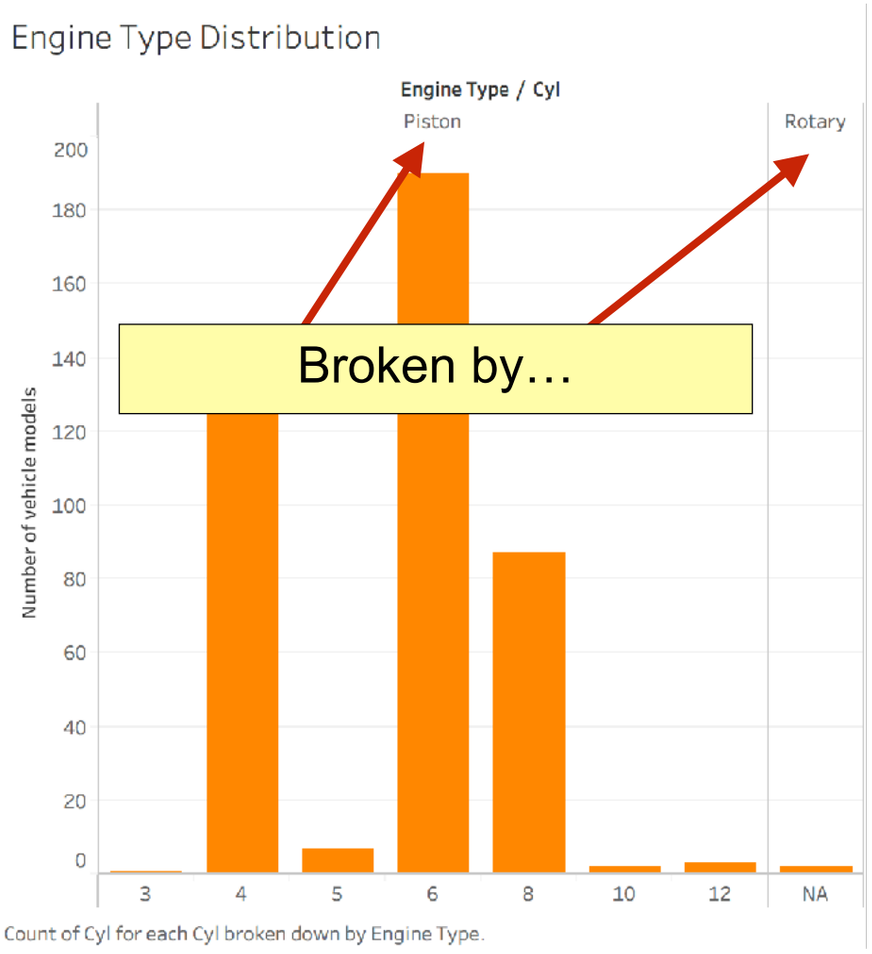


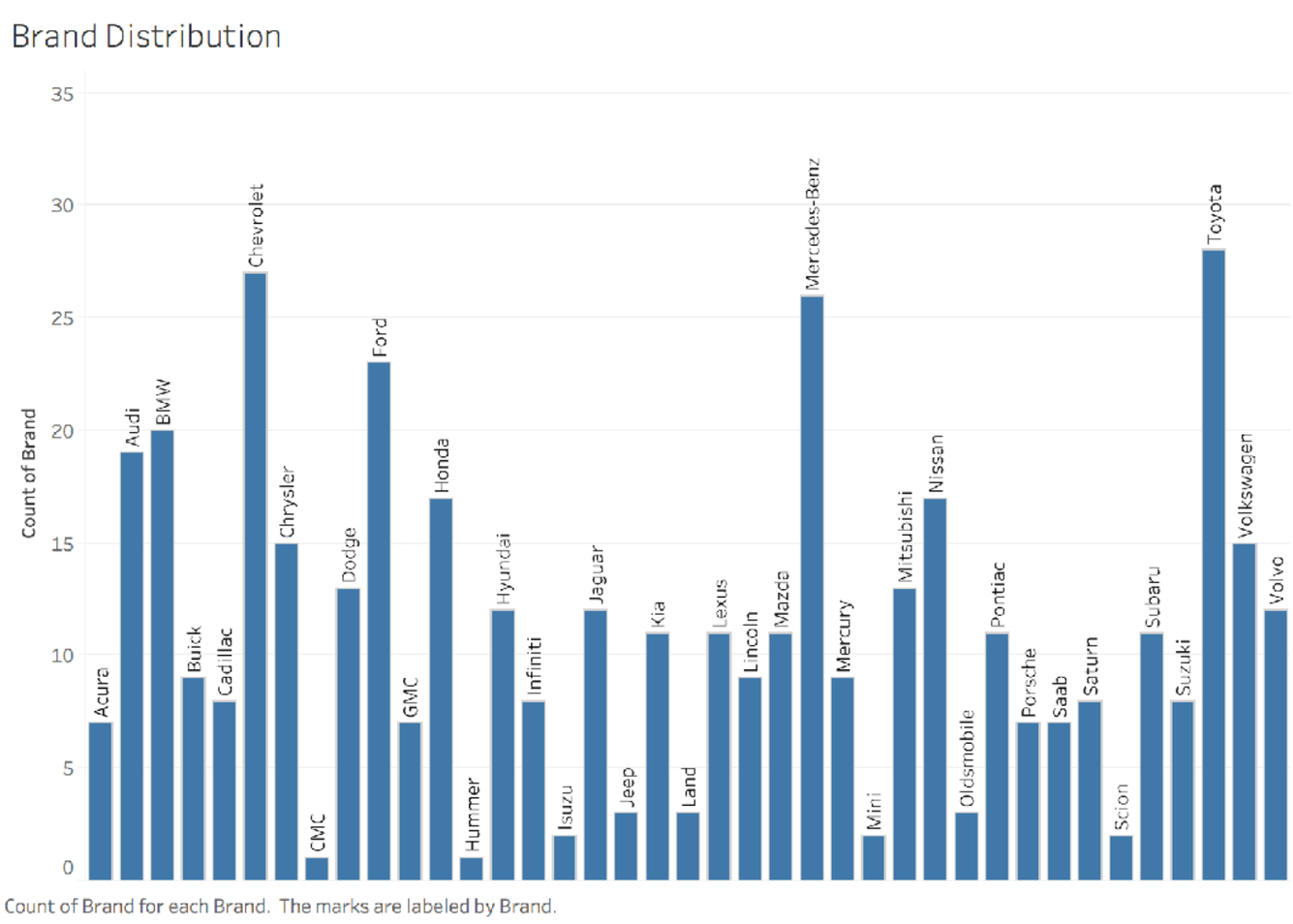


Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Make a **bar chart** of each **discrete variable** counting the occurrences

Additional examples

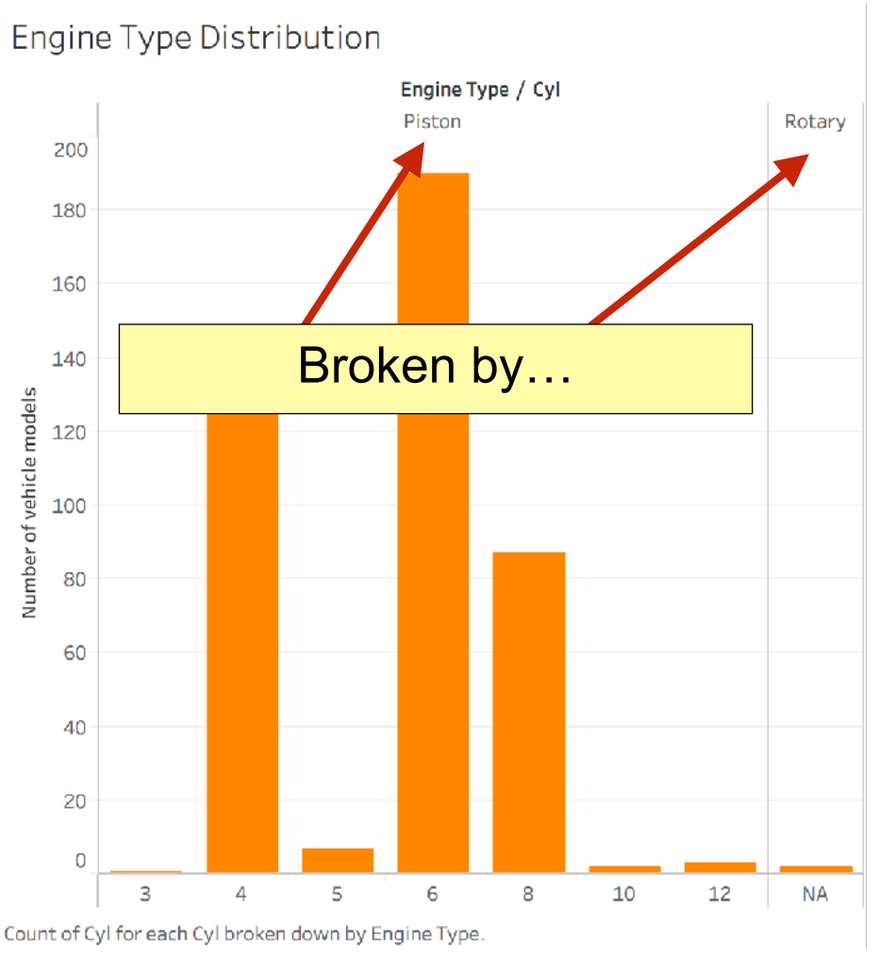


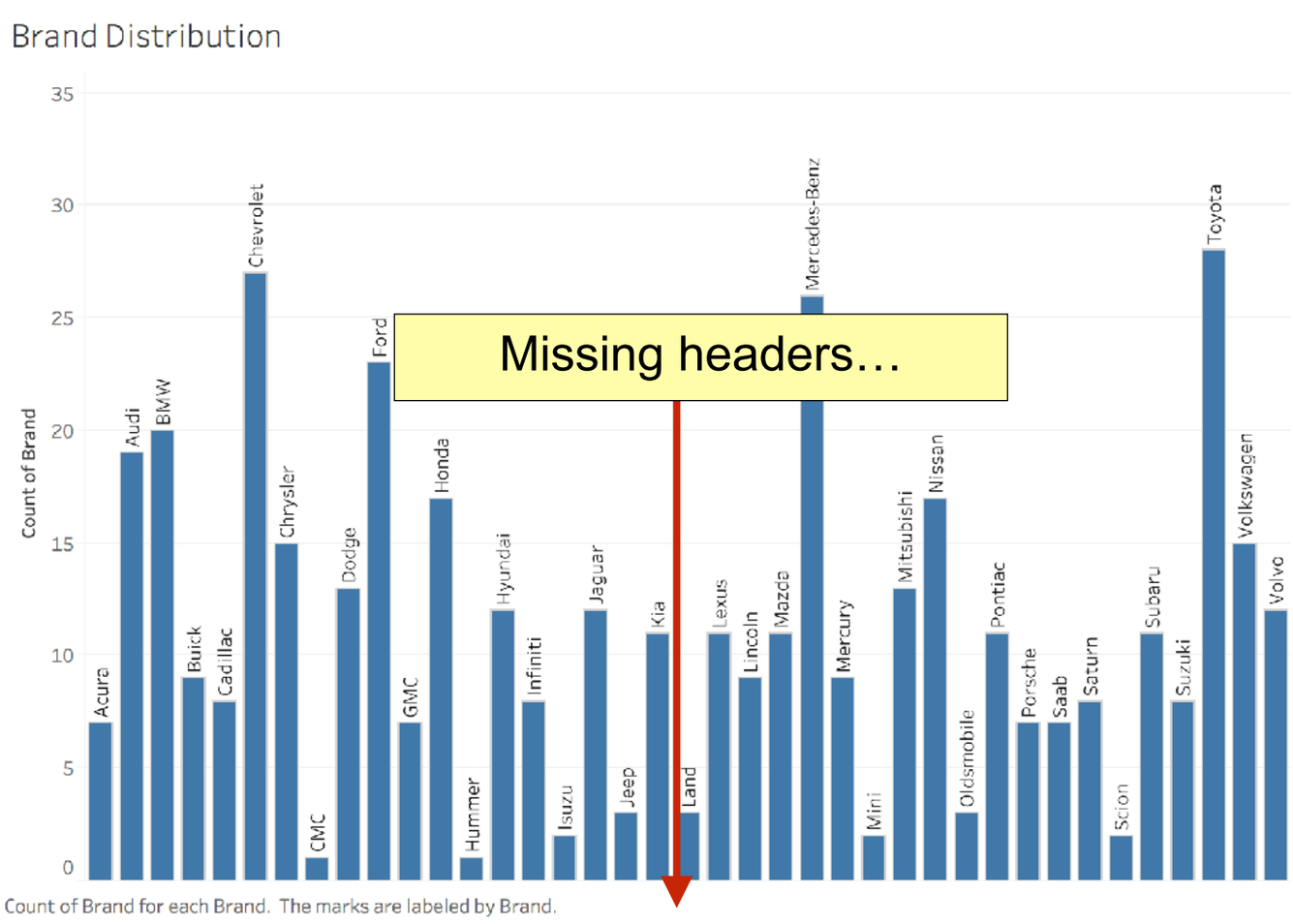


Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Make a **bar chart** of each **discrete variable** counting the occurrences

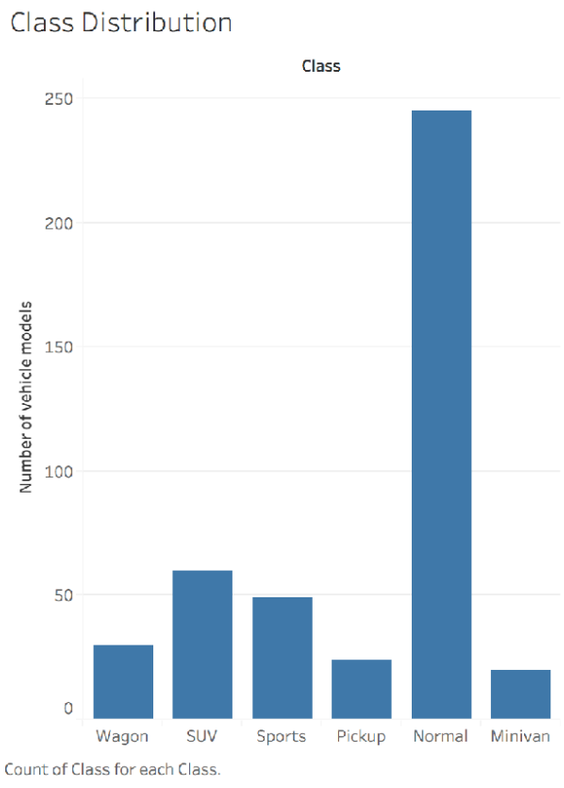
Additional examples

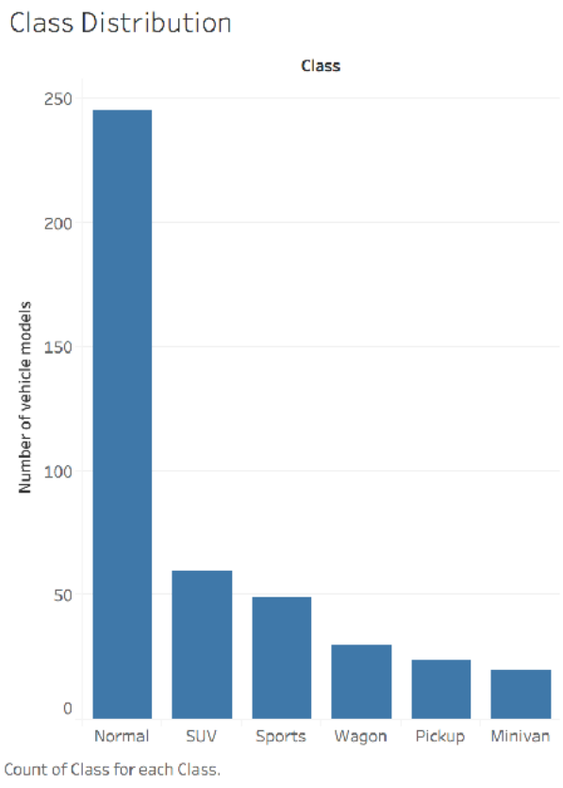


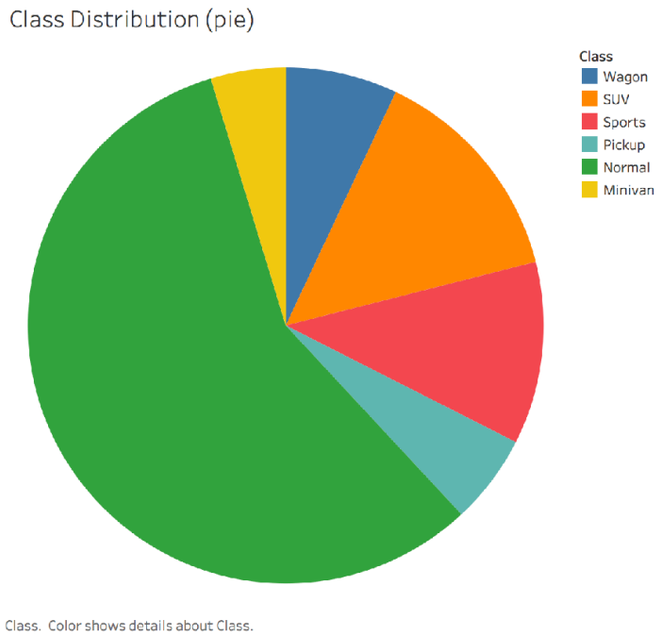


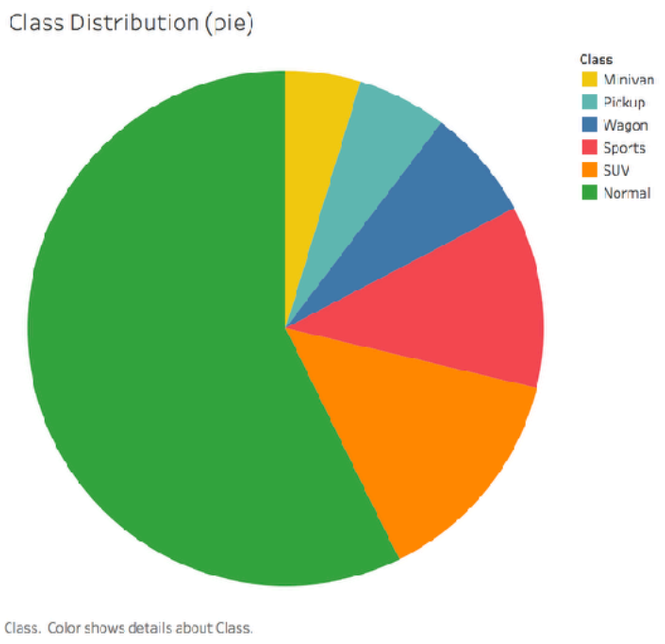
Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Sorting them will help with direct comparisons…



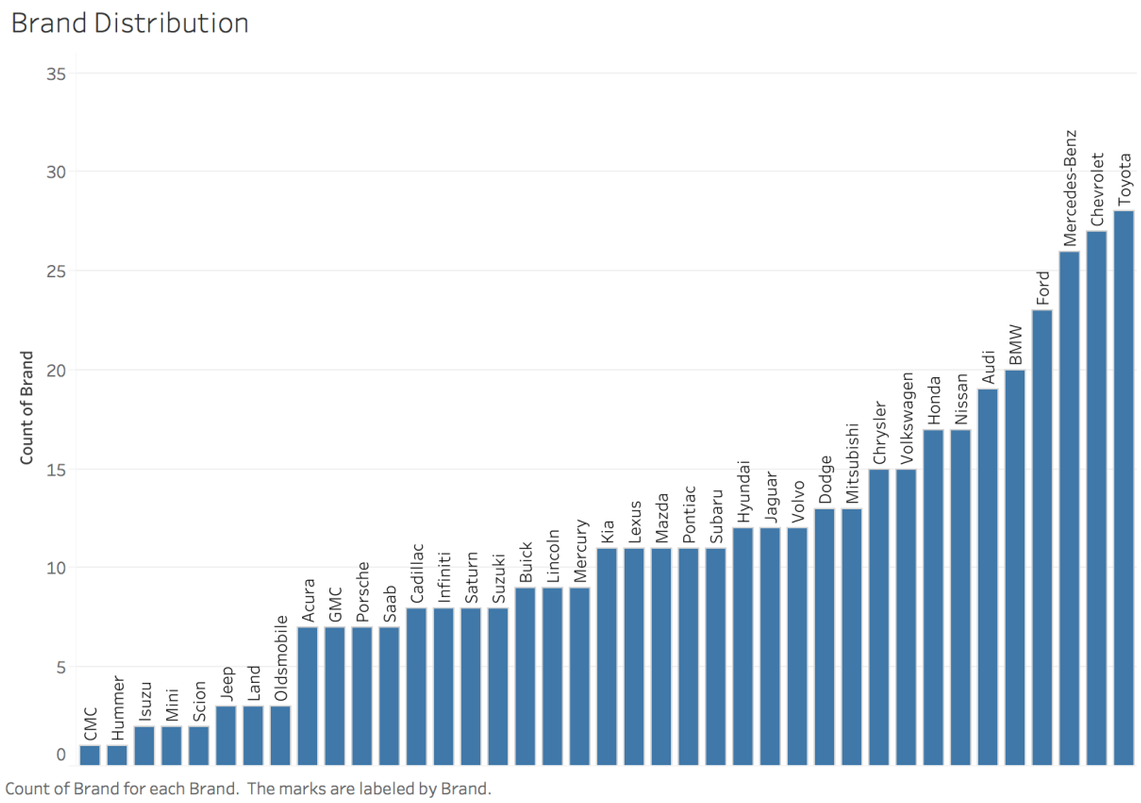


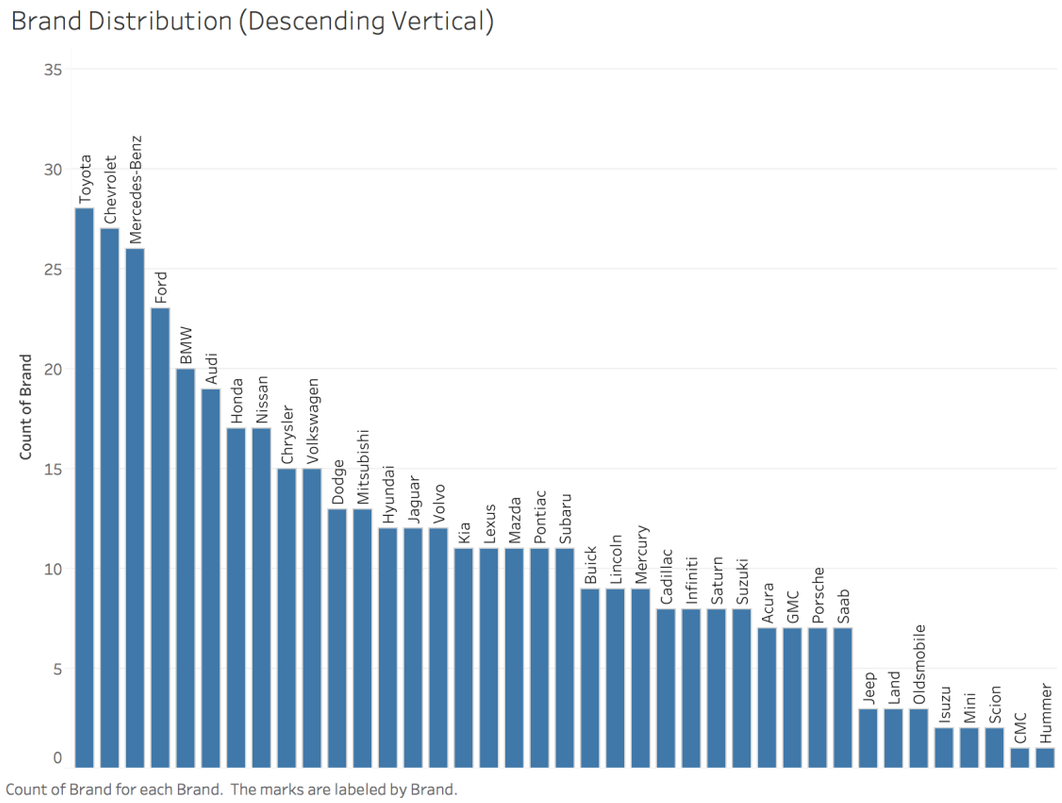




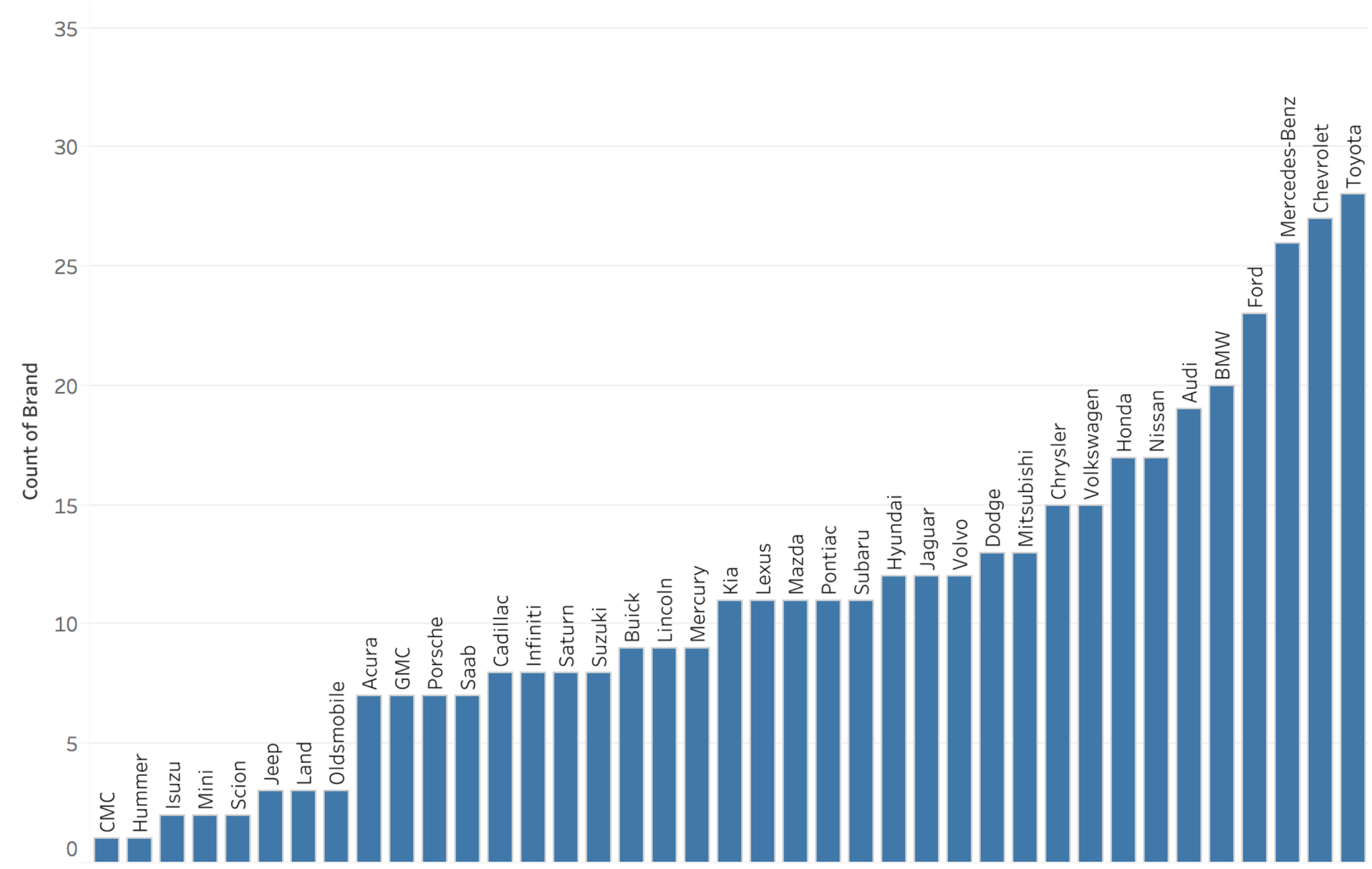
Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Sort Ascending or Sort Descending?
  - Focus in Bottom N or Top N?



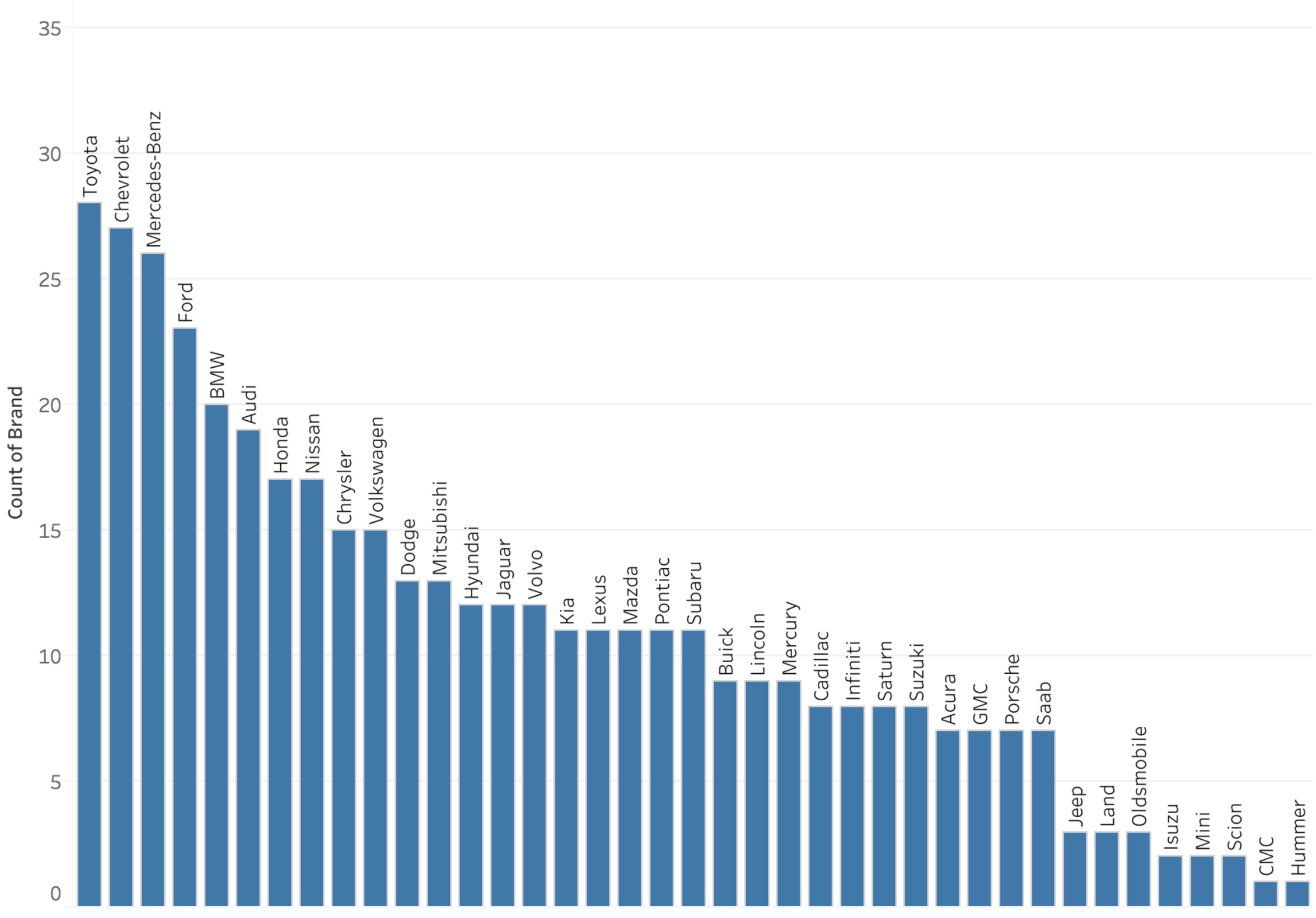


Brand Distribution



Count of Brand for each Brand. The marks are labeled by Brand.

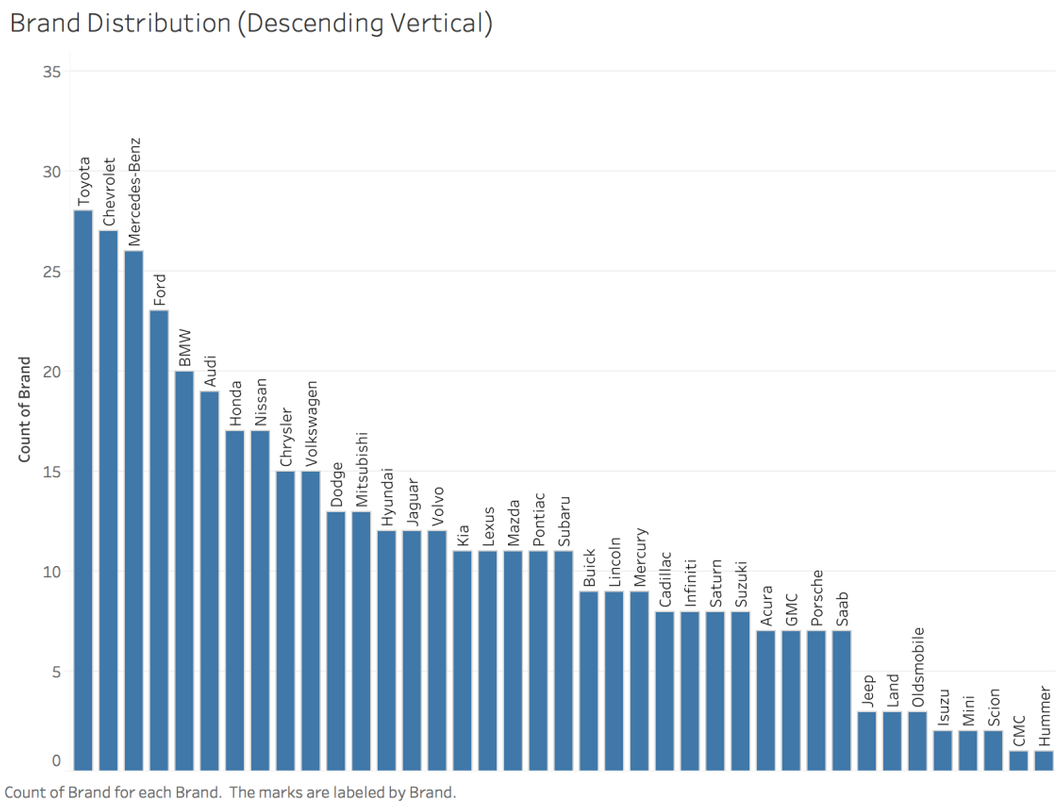
Brand Distribution (Descending Vertical)

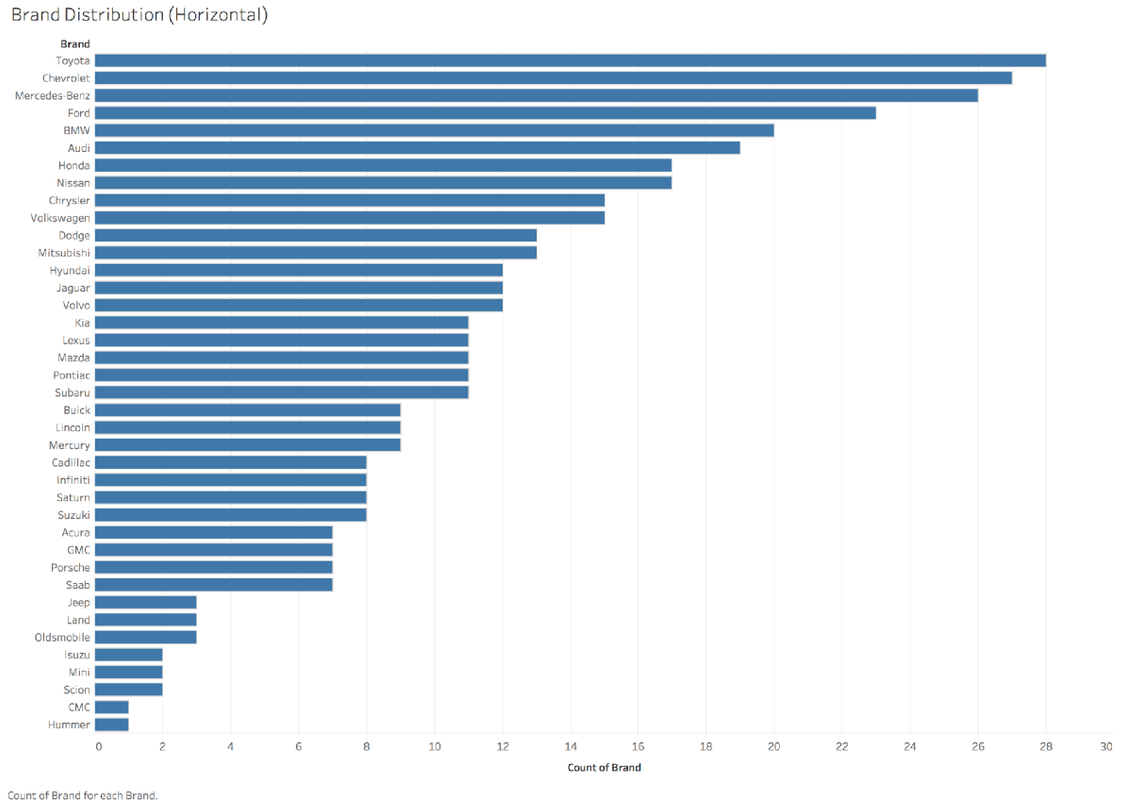


Count of Brand for each Brand. The marks are labeled by Brand.

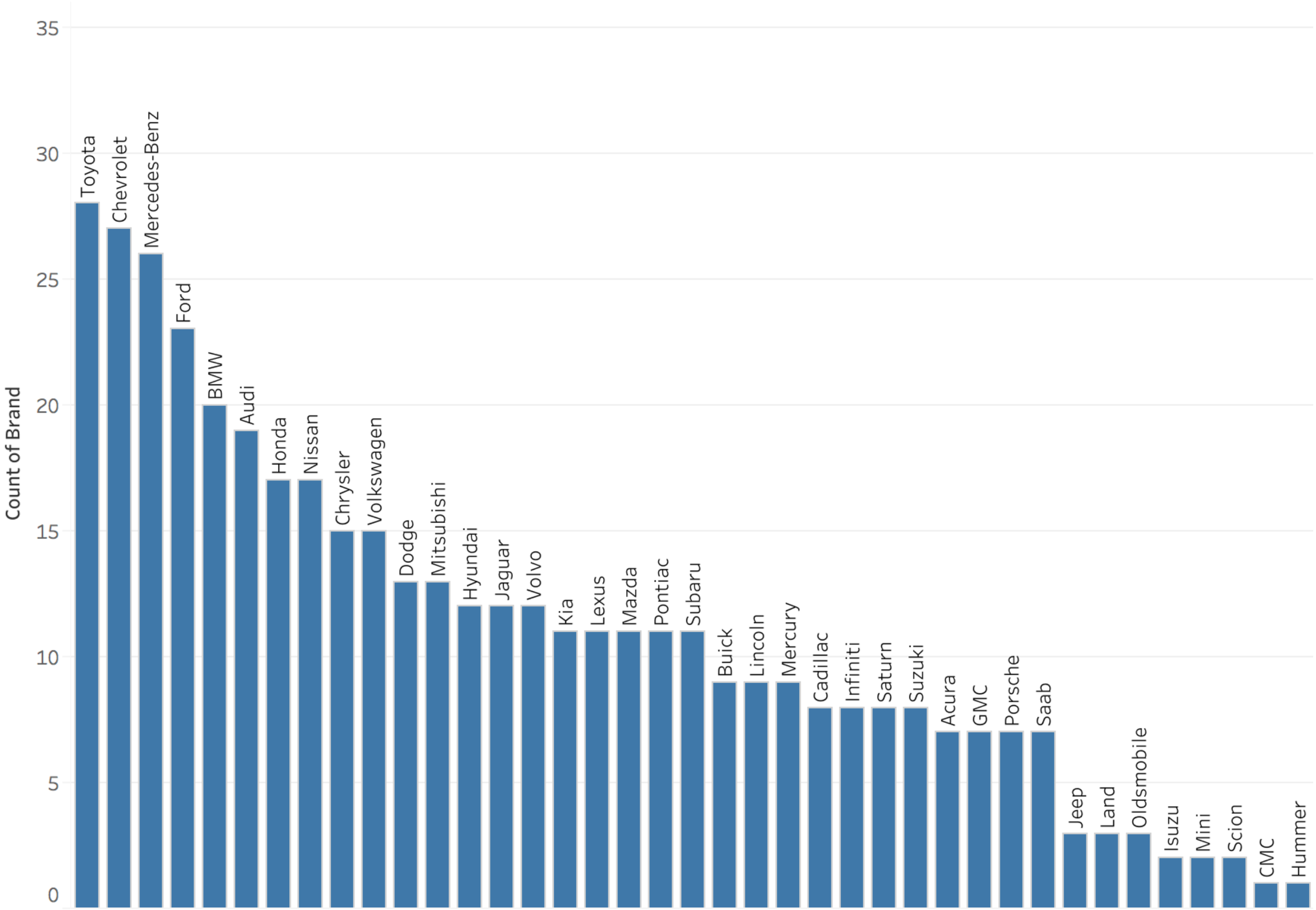
Distribution of a (categorical) variable

- **Compute the distribution of each variable**
  - Vertical or Horizontal layouts?



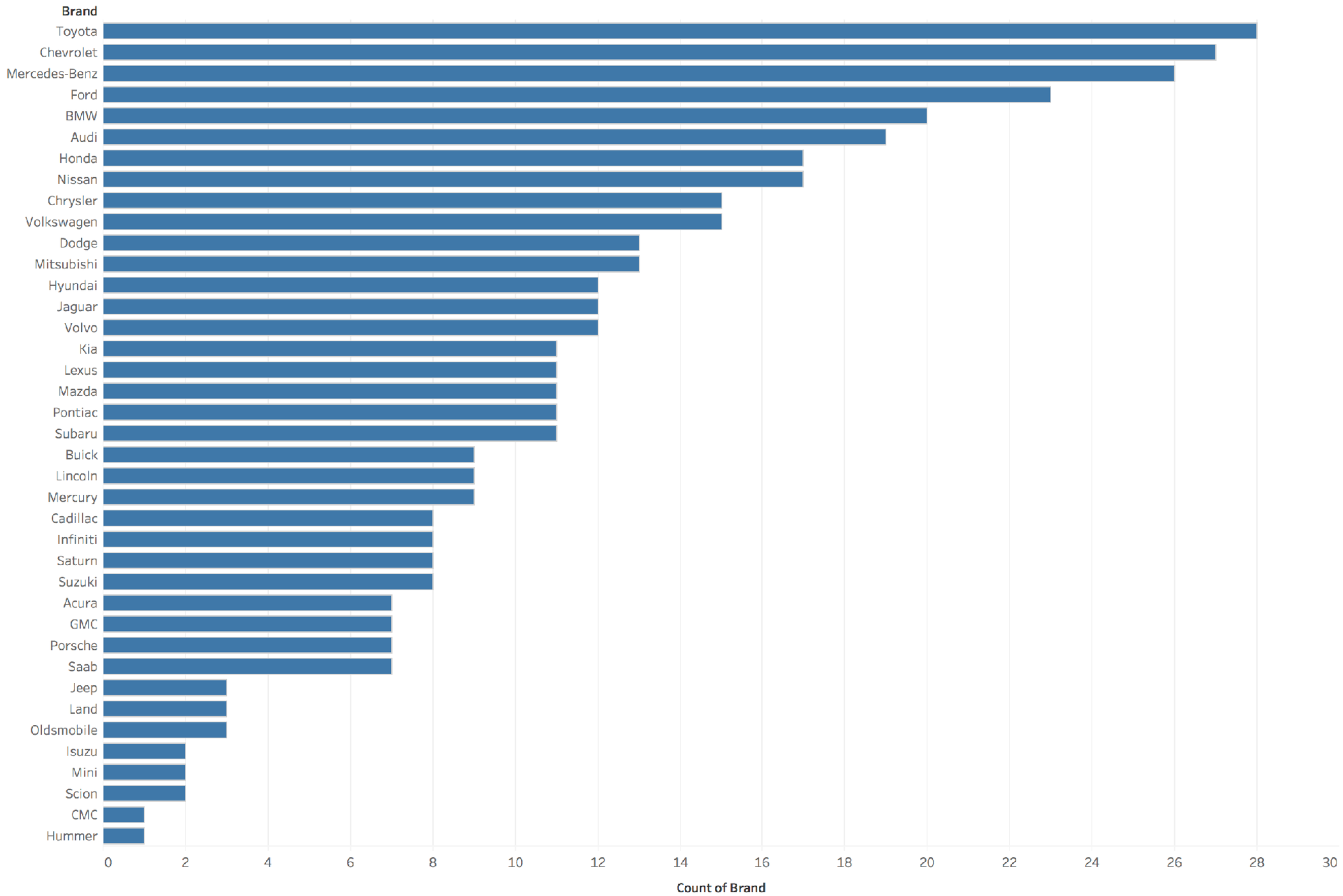


Brand Distribution (Descending Vertical)



Count of Brand for each Brand. The marks are labeled by Brand.

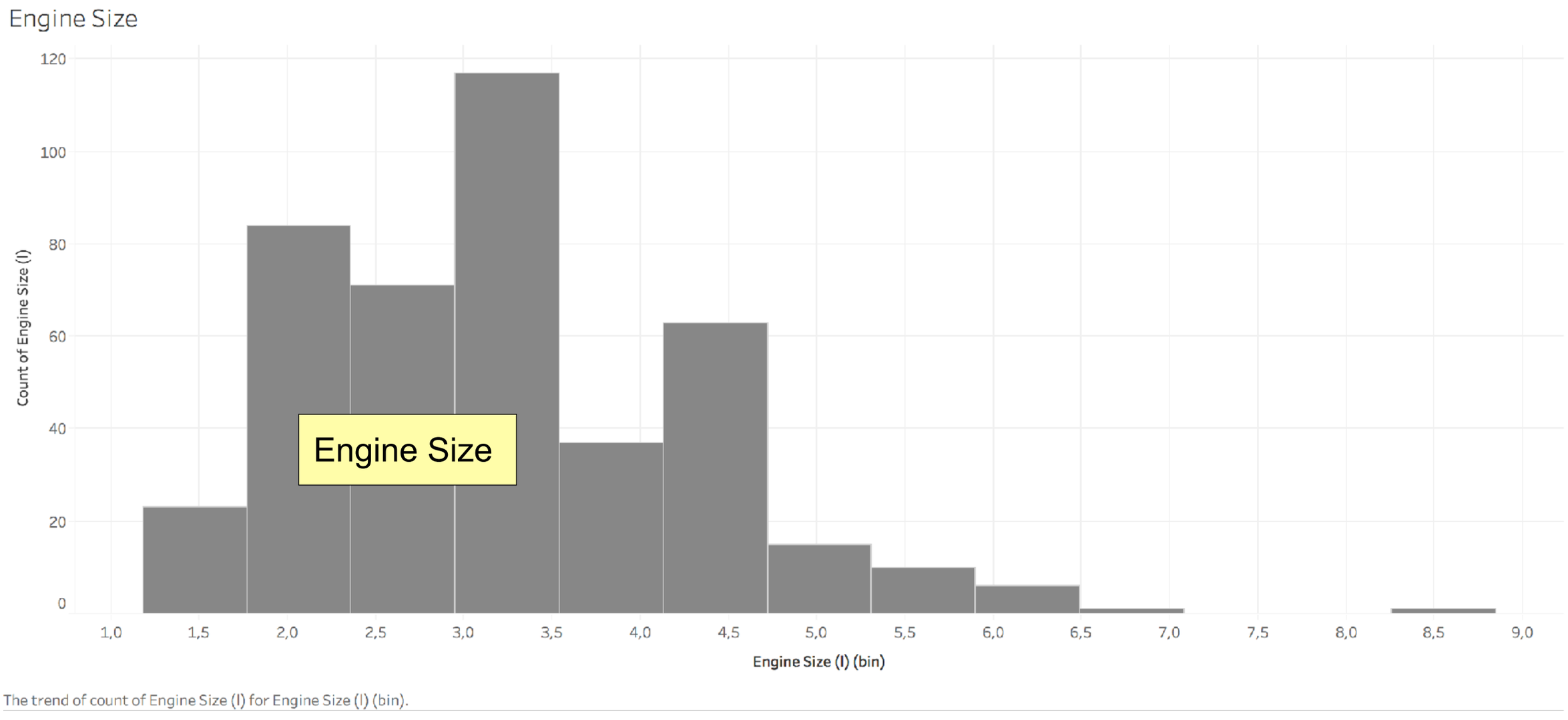
Brand Distribution (Horizontal)



Count of Brand for each Brand.

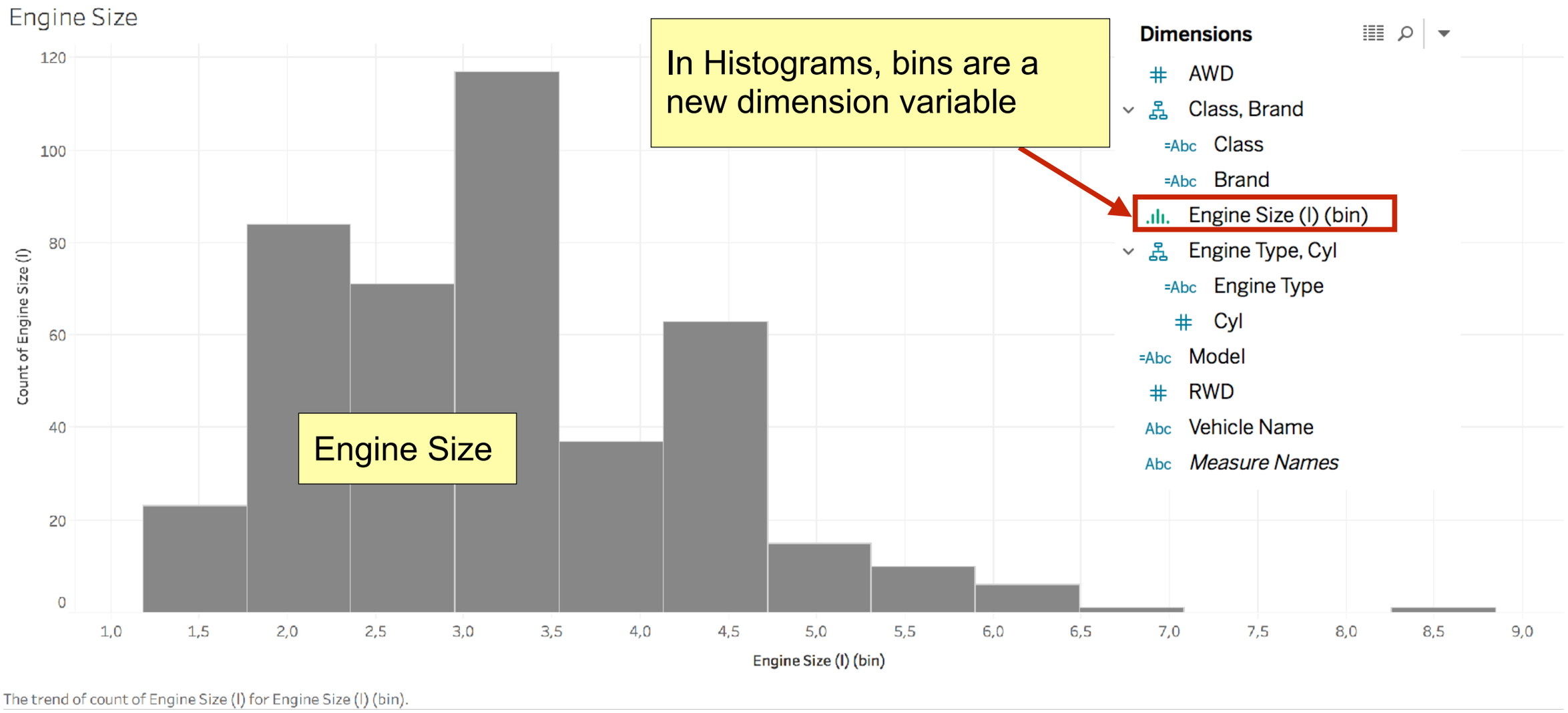
Distribution of a (continuous) variable

- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)



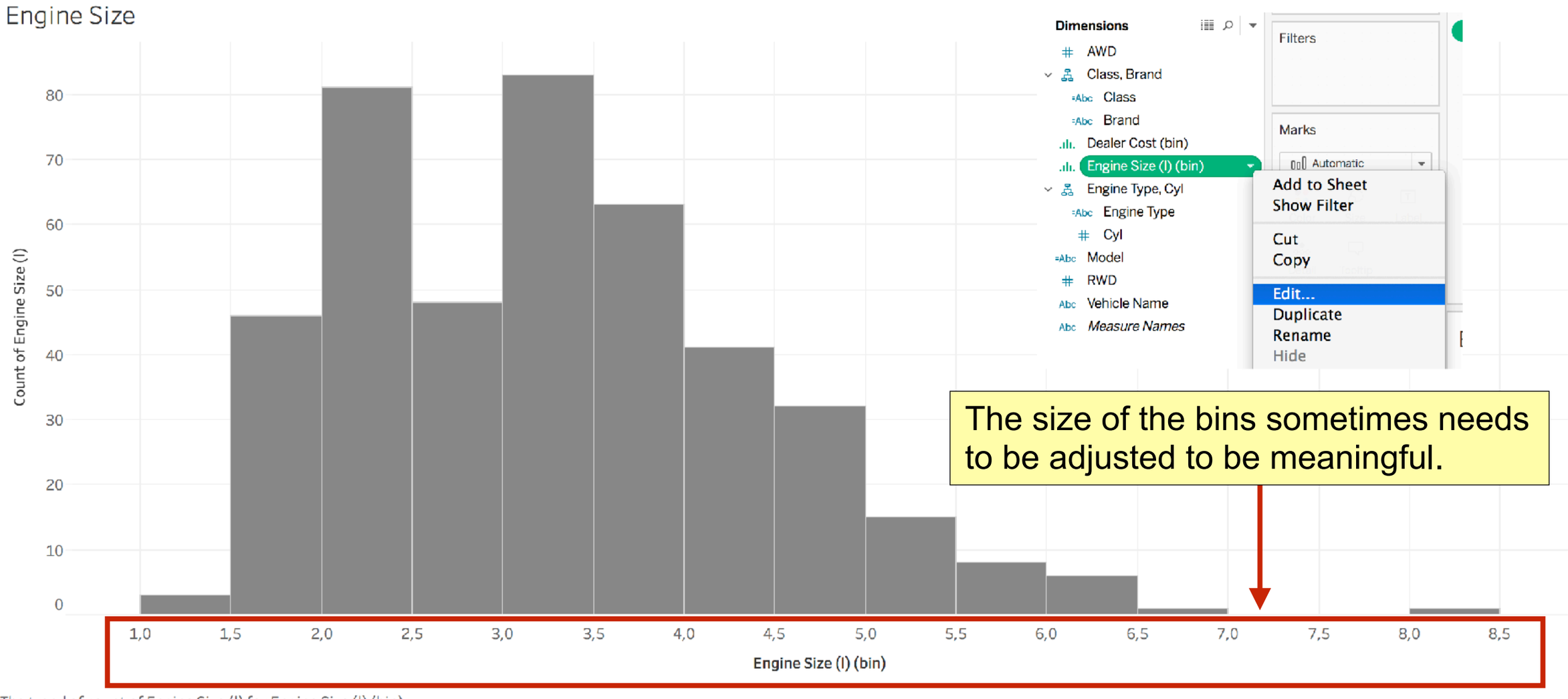
Distribution of a (continuous) variable

- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)



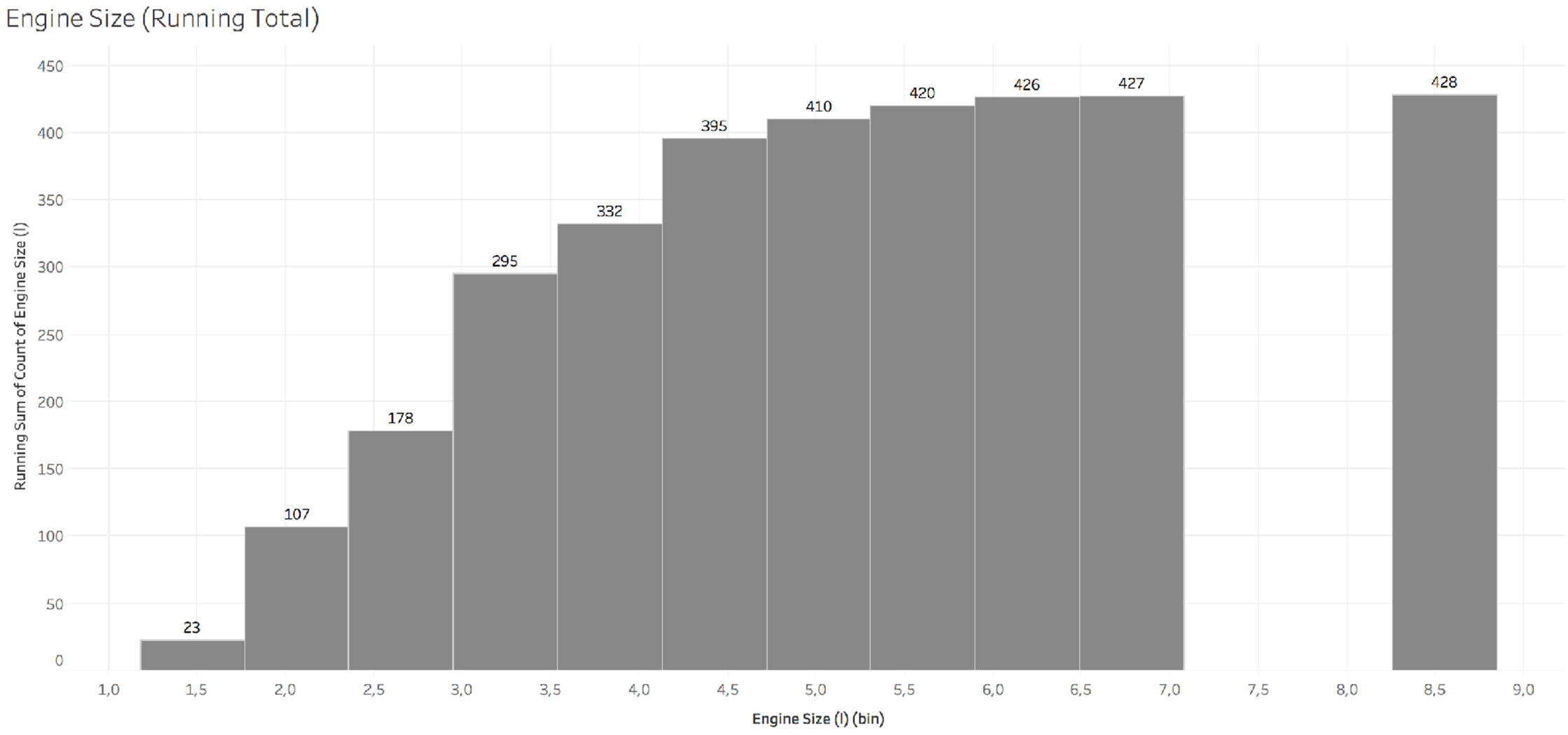
Distribution of a (continuous) variable

- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)



The trend of count of Engine Size (l) for Engine Size (l) (bin).

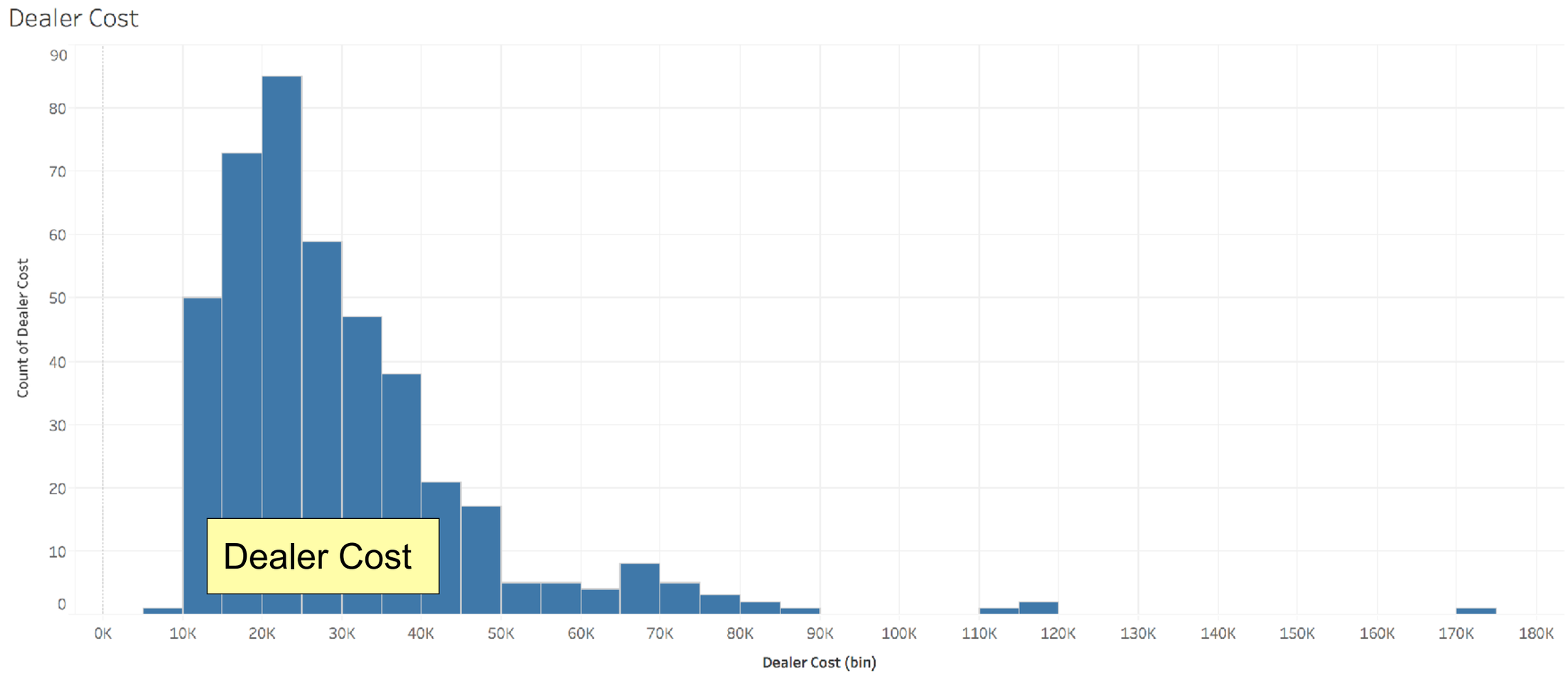
Distribution of a (continuous) variable



The trend of Running Sum of Count of Engine Size (l) for Engine Size (l) (bin).

Distribution of a (continuous) variable

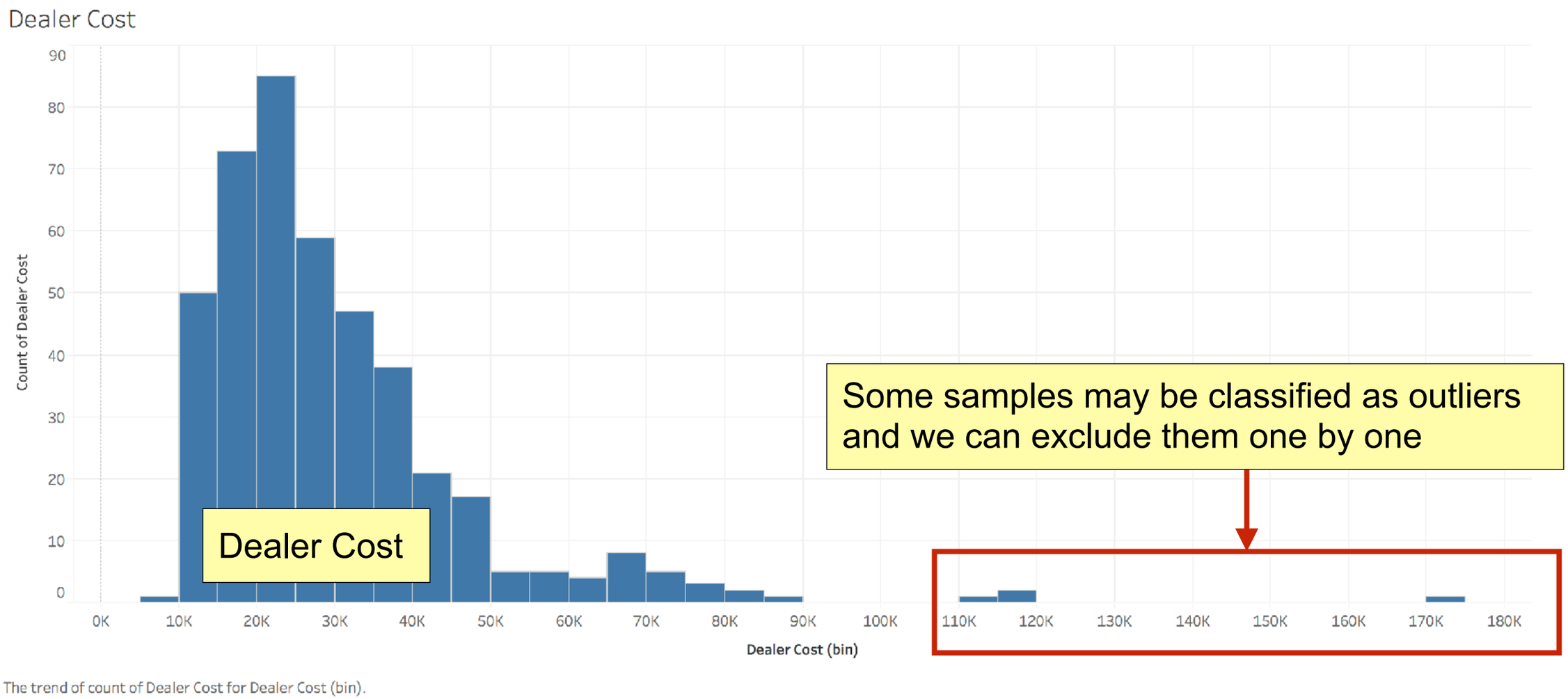
- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)



The trend of count of Dealer Cost for Dealer Cost (bin).

Distribution of a (continuous) variable

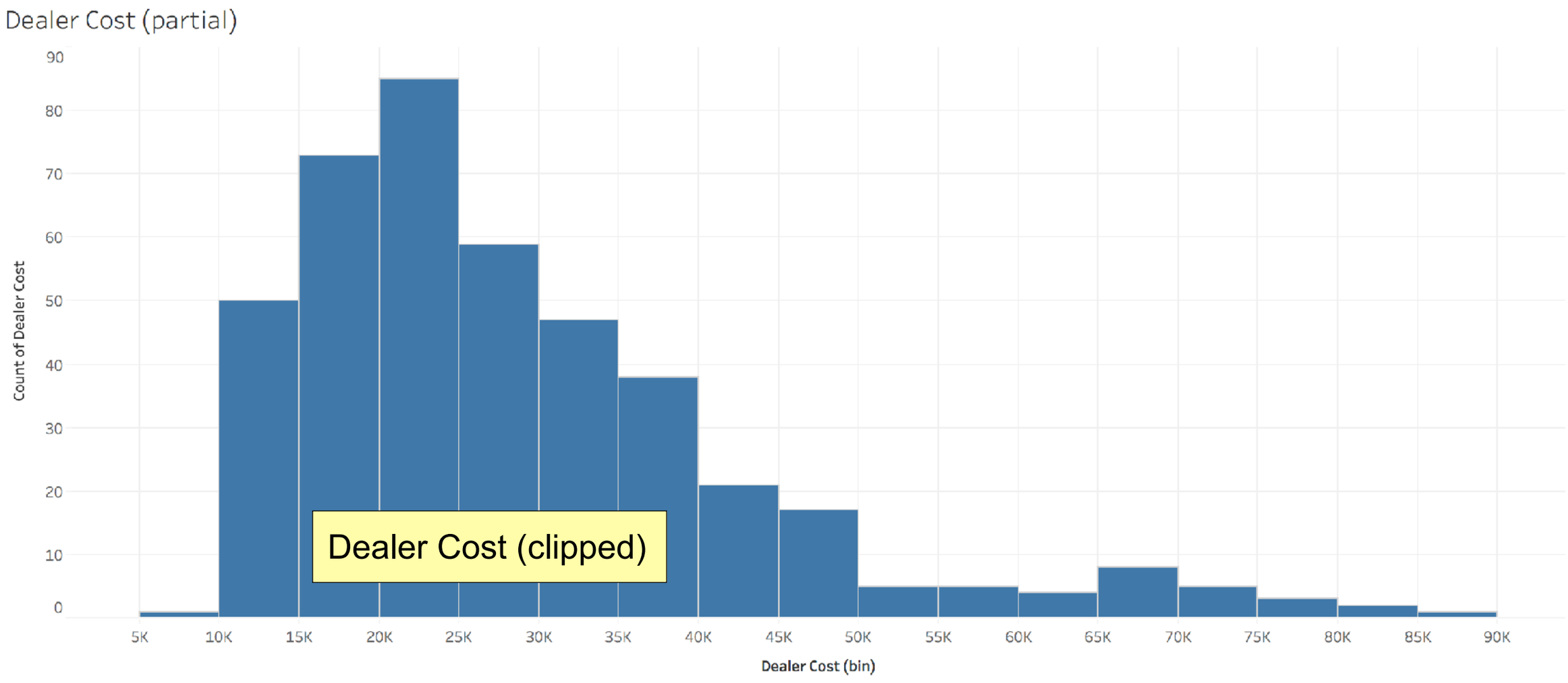
- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)



The trend of count of Dealer Cost for Dealer Cost (bin).

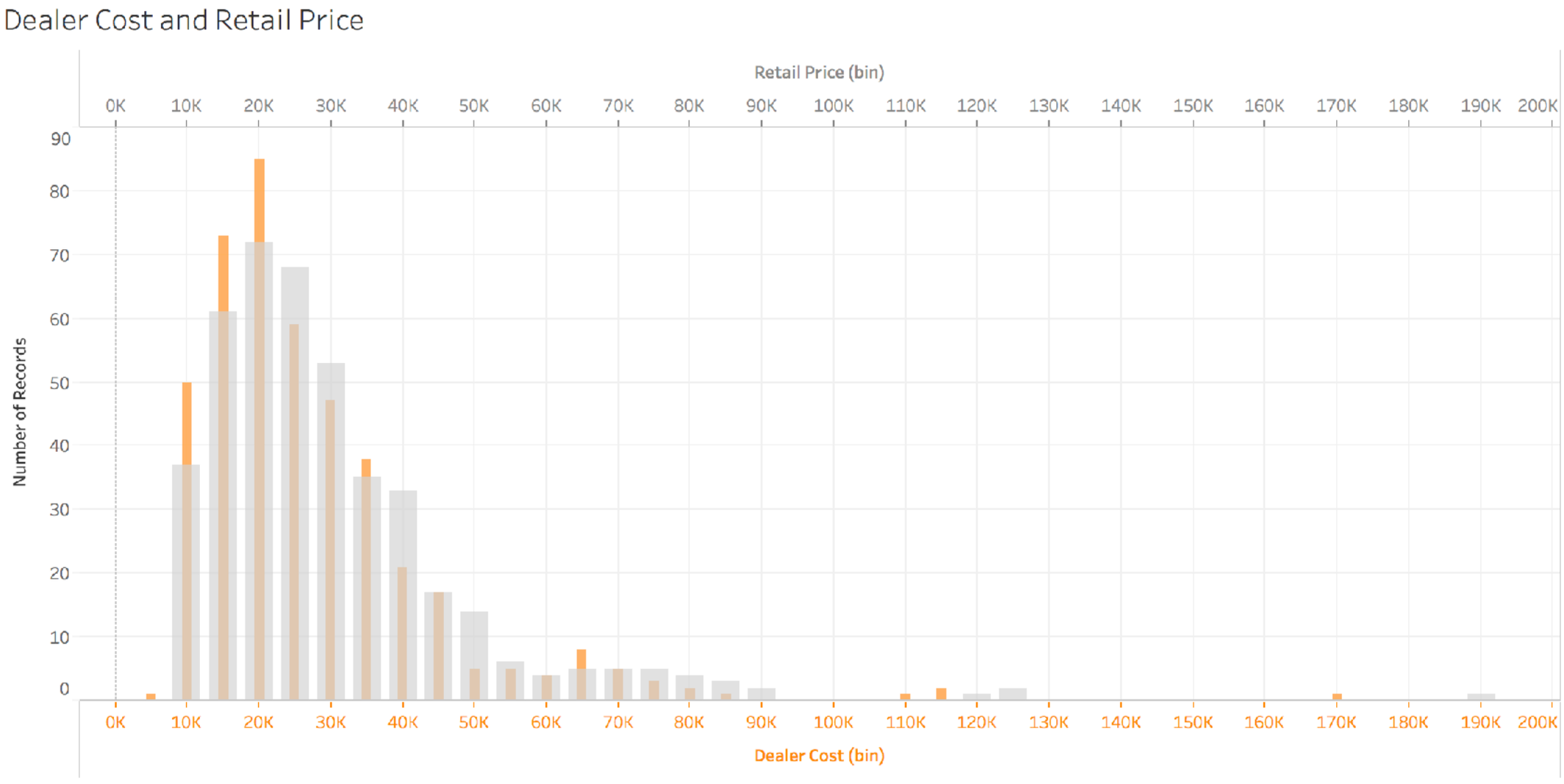
Distribution of a (continuous) variable

- **Compute the distribution of each variable**
  - Make an **histogram** of each **continuous variable** by creating bins (intervals)

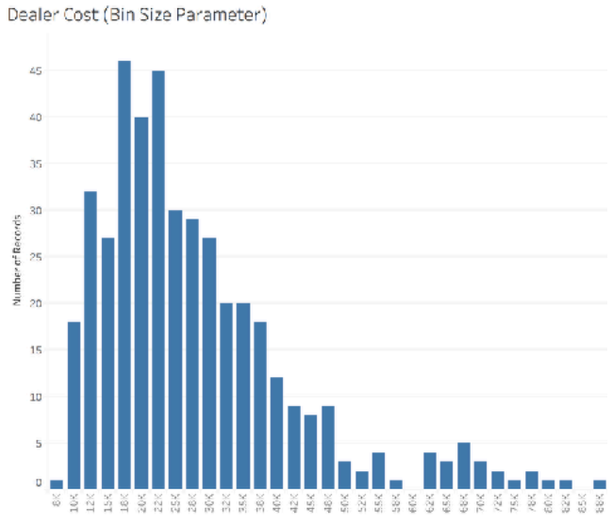


The trend of count of Dealer Cost for Dealer Cost (bin). The data is filtered on Dealer Cost (bin), which excludes 110K, 115K and 170K.

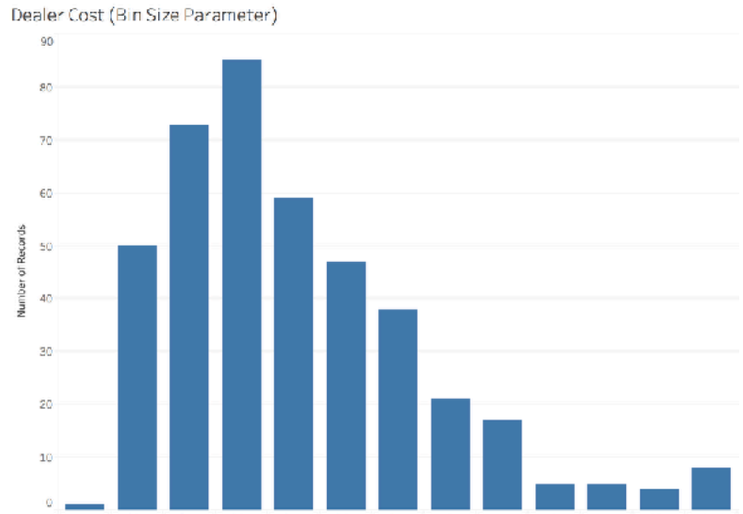
Distribution of a (continuous) variable



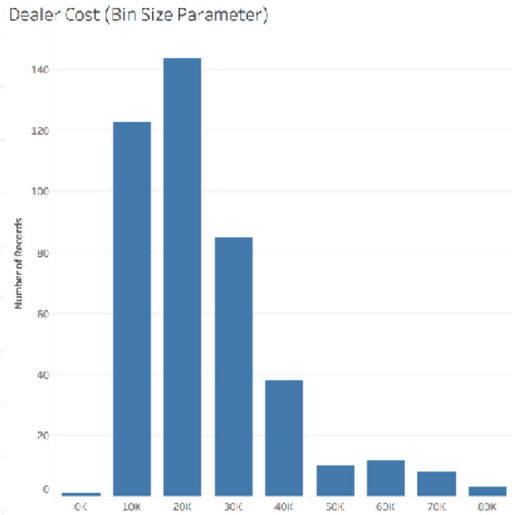
Distribution of a (continuous) variable



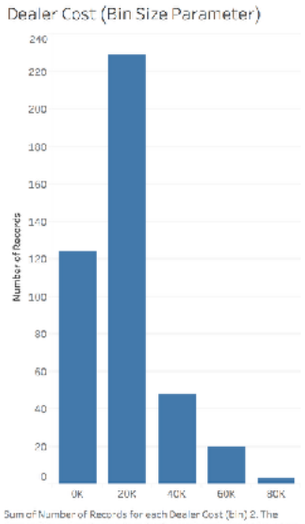
Sum of Number of Records for each Dealer Cost (bin) 2. The data is filtered on Dealer Cost (bin), which excludes 110K, 115K and 170K.



Sum of Number of Records for each Dealer Cost (bin) 2. The data is filtered on Dealer Cost (bin), which excludes 110K, 115K and 170K.

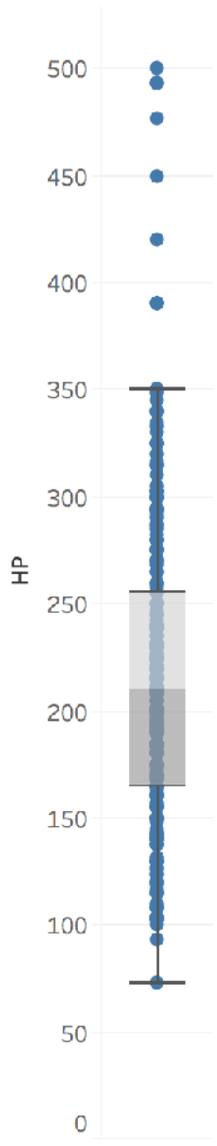


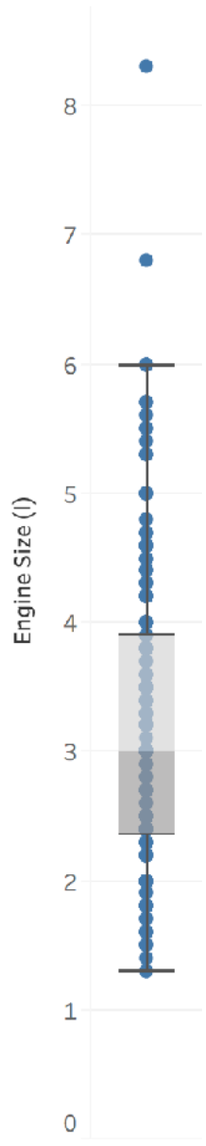
Sum of Number of Records for each Dealer Cost (bin) 2. The data is filtered on Dealer Cost (bin), which excludes 110K, 115K and 170K.

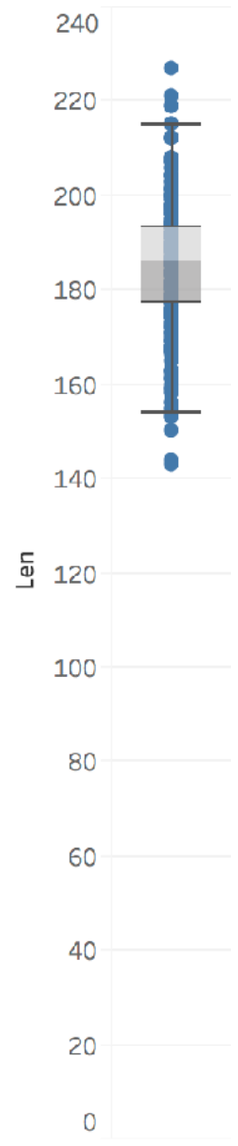


Sum of Number of Records for each Dealer Cost (bin) 2. The data is filtered on Dealer Cost (bin), which excludes 110K, 115K and 170K.

Distribution of a (continuous) variable

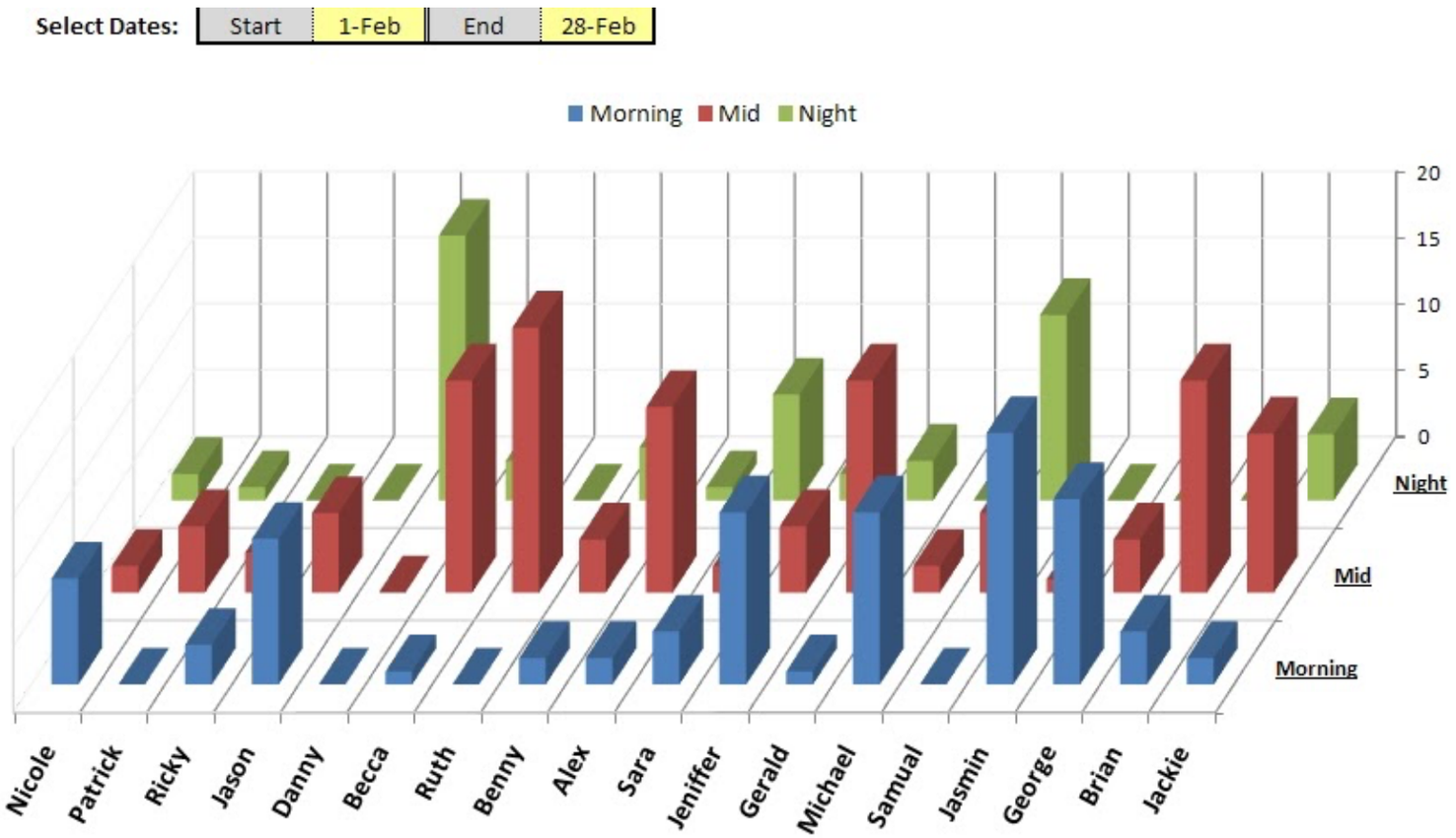






Explore the Data: Two Variables

Joint Distribution of (2 categorical) variables



Joint Distribution of (2 categorical) variables

- **Create a double entry table and map the number of occurrences to color**
- **The heat map created has rectangular cells**

	Class					
Brand	Minivan	Normal	Pickup	Sports	SUV	Wagon
Acura		5		1	1	
Audi		13		4		2
BMW		13		4	2	1
Buick		7			2	
Cadillac		4	1	1	2	
Chevrolet	2	13	5	2	4	1
Chrysler	2	11		1		1
CMC					1	
Dodge	2	6	3	1	1	
Ford	1	10	3	3	4	2
GMC	1		4		2	
Honda	2	11		1	3	
Hummer					1	
Hyundai		10		1	1	
Infiniti		6				2
Isuzu					2	
Jaguar		8		4		
Jeep					3	
Kia	1	8			1	1
Land					3	
Lexus		6		1	3	1
Lincoln		7			2	
Mazda	1	3	2	4	1	
Mercedes-Benz		16		5	2	3
Mercury	1	6			1	1
Mini		2				
Mitsubishi		6		3	3	1
Nissan	2	7	2	2	3	1
Oldsmobile	1	3				

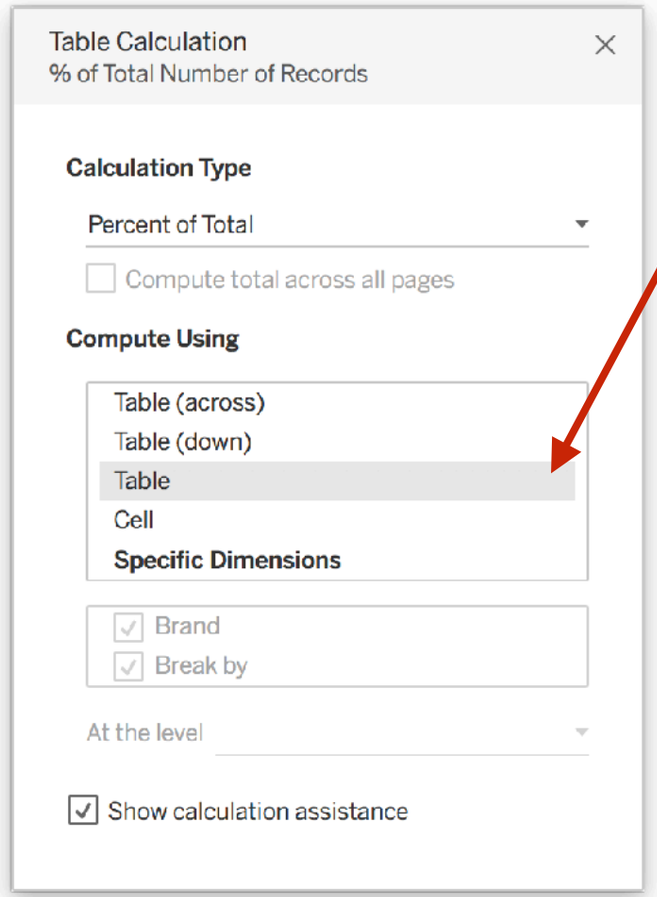
	Break by		
Brand	AWD	FWD	RWD
Acura	1	5	1
Audi	12	7	
BMW	5		15
Buick	1	8	
Cadillac	1	5	2
Chevrolet	5	18	4
Chrysler		13	2
CMC		1	
Dodge	2	7	4
Ford	4	12	7
GMC	3	1	3
Honda	3	13	1
Hummer	1		
Hyundai		12	
Infiniti	2	1	5
Isuzu	1	1	
Jaguar	2		10
Jeep	2	1	
Kia		11	
Land	3		
Lexus	3	1	7
Lincoln	1	1	7
Mazda	2	4	5
Mercedes-Benz	6		20
Mercury		5	4
Mini		2	
Mitsubishi	2	11	
Nissan	2	12	3
Oldsmobile		3	

	Break by							
Brand	3	4	5	6	8	10	12	-1
Acura		2		5				
Audi		4		10	5			
BMW				16	4			
Buick				9				
Cadillac				1	7			
Chevrolet		7		13	7			
Chrysler		5		10				
CMC					1			
Dodge		5		6	1	1		
Ford		7		7	8	1		
GMC		1		3	3			
Honda	1	11		5				
Hummer					1			
Hyundai		6		6				
Infiniti				5	3			
Isuzu				2				
Jaguar				3	9			
Jeep		1		2				
Kia		7		4				
Land				1	2			
Lexus				6	5			
Lincoln				2	7			
Mazda		7		2				2
Mercedes-Benz		2		12	10		2	
Mercury				5	4			
Mini		2						
Mitsubishi		7		6				
Nissan		4		11	2			
Oldsmobile		1		2				

Joint Distribution of (2 categorical) variables

Brand	Break by: Minivan	Normal	Pickup	Sports	SUV	Wagon
Acura		1,168% [1]		0,234% [2]	0,234% [3]	
Audi		3,037% [4]		0,935% [5]		0,467% [6]
BMW		3,037% [7]		0,935% [8]	0,467% [9]	0,234% [10]
Buick		1,636% [11]			0,467% [12]	
Cadillac		0,935% [13]	0,234% [14]	0,234% [15]	0,467% [16]	
Chevrolet	0,467% [17]	3,037% [18]	1,168% [19]	0,467% [20]	0,935% [21]	0,234% [22]
Chrysler	0,467% [23]	2,570% [24]		0,234% [25]		0,234% [26]
CMC					0,234% [27]	
Dodge	0,467% [28]	1,402% [29]	0,701% [30]	0,234% [31]	0,234% [32]	
Ford	0,234% [33]	2,336% [34]	0,701% [35]	0,701% [36]	0,935% [37]	0,467% [38]
GMC	0,234% [39]		0,935% [40]		0,467% [41]	
Honda	0,467% [42]	2,570% [43]		0,234% [44]	0,701% [45]	
Hummer					0,234% [46]	
Hyundai		2,336% [47]		0,234% [48]	0,234% [49]	
Infiniti		1,402% [50]				0,467% [51]
Isuzu					0,467% [52]	
		1,869%		0,935%		

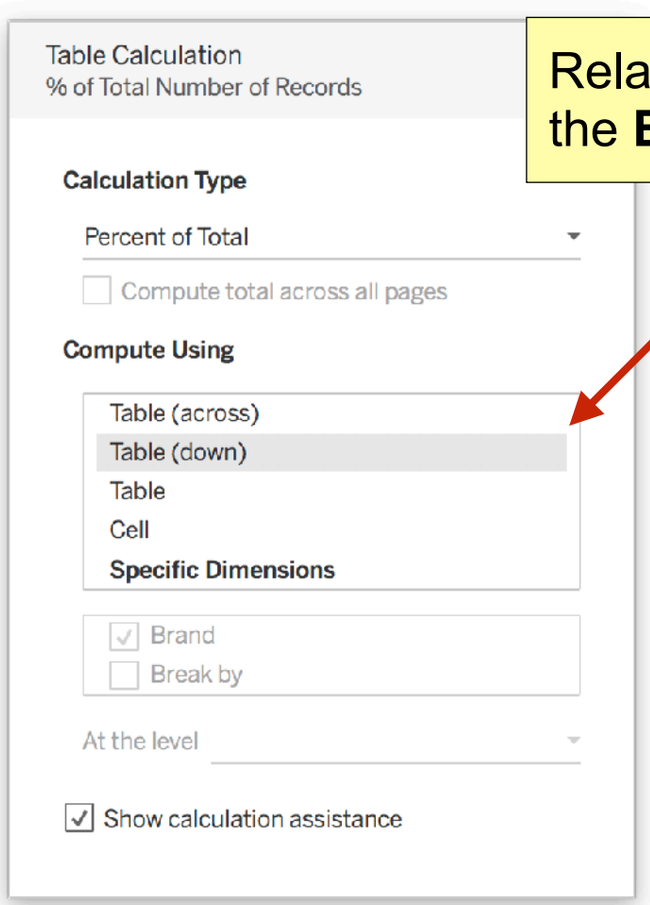
Here the Percent of Total is relative to the whole table using



Joint Distribution of (2 categorical) variables

Break by

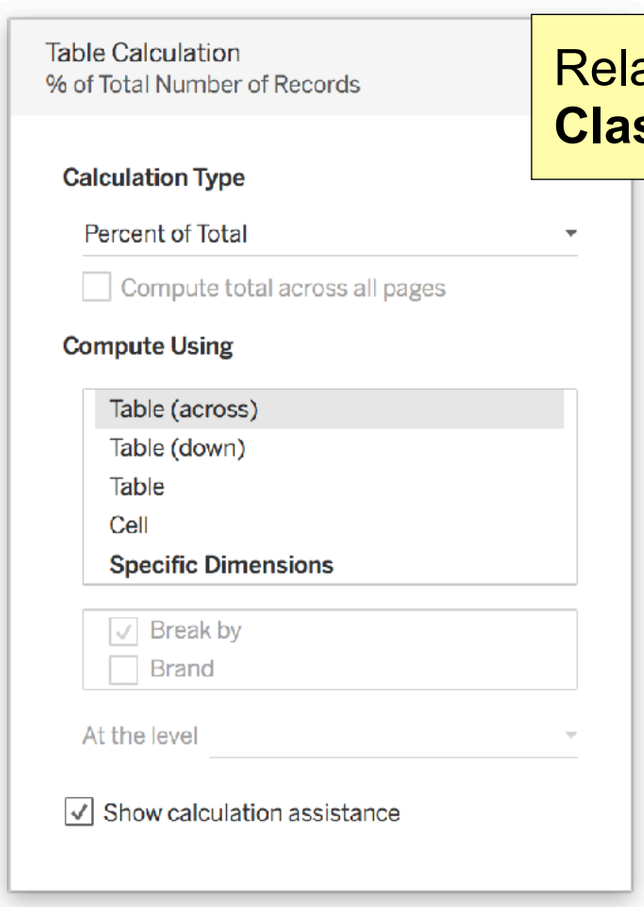
Brand	Minivan	Normal	Pickup	Sports	SUV	Wagon
Acura	[1]	2,04% [1]	[1]	2,04% [1]	1,67% [1]	[1]
Audi	[2]	5,31% [2]	[2]	8,16% [2]	[2]	6,67% [2]
BMW	[3]	5,31% [3]	[3]	8,16% [3]	3,33% [3]	3,33% [3]
Buick	[4]	2,86% [4]	[4]	[4]	3,33% [4]	[4]
Cadillac	[5]	1,63% [5]	4,17% [5]	2,04% [5]	3,33% [5]	[5]
Chevrolet	10,00% [6]	5,31% [6]	20,83% [6]	4,08% [6]	6,67% [6]	3,33% [6]
Chrysler	10,00% [7]	4,49% [7]	[7]	2,04% [7]	[7]	3,33% [7]
CMC	[8]	[8]	[8]	[8]	1,67% [8]	[8]
Dodge	10,00% [9]	2,45% [9]	12,50% [9]	2,04% [9]	1,67% [9]	[9]
Ford	5,00% [10]	4,08% [10]	12,50% [10]	6,12% [10]	6,67% [10]	6,67% [10]
GMC	5,00% [11]	[11]	16,67% [11]	[11]	3,33% [11]	[11]
Honda	10,00% [12]	4,49% [12]	[12]	2,04% [12]	5,00% [12]	[12]
Hummer	[13]	[13]	[13]	[13]	1,67% [13]	[13]
Hyundai	[14]	4,08% [14]	[14]	2,04% [14]	1,67% [14]	[14]
Infiniti	[15]	2,45% [15]	[15]	[15]	[15]	6,67% [15]
Isuzu	[16]	[16]	[16]	[16]	3,33% [16]	[16]
Jaguar	[17]	3,27%	[17]	8,16%	[17]	[17]



Relative to a column (which is exactly the **Brand** dimension in this example)

Joint Distribution of (2 categorical) variables

Brand	Break by: Minivan	Normal	Pickup	Sports	SUV	Wagon
Acura	[1]	71,43% [2]	[3]	14,29% [4]	14,29% [5]	[6]
Audi	[1]	68,42% [2]	[3]	21,05% [4]	[5]	10,53% [6]
BMW	[1]	65,00% [2]	[3]	20,00% [4]	10,00% [5]	5,00% [6]
Buick	[1]	77,78% [2]	[3]	[4]	22,22% [5]	[6]
Cadillac	[1]	50,00% [2]	12,50% [3]	12,50% [4]	25,00% [5]	[6]
Chevrolet	7,41% [1]	48,15% [2]	18,52% [3]	7,41% [4]	14,81% [5]	3,70% [6]
Chrysler	13,33% [1]	73,33% [2]	[3]	6,67% [4]	[5]	6,67% [6]
CMC	[1]	[2]	[3]	[4]	100,00% [5]	[6]
Dodge	15,38% [1]	46,15% [2]	23,08% [3]	7,69% [4]	7,69% [5]	[6]
Ford	4,35% [1]	43,48% [2]	13,04% [3]	13,04% [4]	17,39% [5]	8,70% [6]
GMC	14,29% [1]	[2]	57,14% [3]	[4]	28,57% [5]	[6]
Honda	11,76% [1]	64,71% [2]	[3]	5,88% [4]	17,65% [5]	[6]
Hummer	[1]	[2]	[3]	[4]	100,00% [5]	[6]
Hyundai	[1]	83,33% [2]	[3]	8,33% [4]	8,33% [5]	[6]
Infiniti	[1]	75,00% [2]	[3]	[4]	[5]	25,00% [6]
Isuzu	[1]	[2]	[3]	[4]	100,00% [5]	[6]
Jaguar	[1]	66,67% [2]	[3]	33,33% [4]	[5]	[6]



Relative to a row (which is exactly the **Class** dimension in this example)

Joint Distribution of (2 categorical) variables

- **Exercise…**

Class

Brand	Minivan	Normal	Pickup	Sports	SUV	Wagon
Acura		5		1	1	
Audi		13		4		2
BMW		13		4	2	1
Buick		7			2	
Cadillac		4	1	1	2	
Chevrolet	2	13	5	2	4	1
Chrysler	2	11		1		1
CMC					1	
Dodge	2	6	3	1	1	
Ford	1	10	3	3	4	2
GMC	1		4		2	
Honda	2	11		1	3	
Hummer					1	
Hyundai		10		1	1	
Infiniti		6				
Isuzu					2	
Jaguar		8		4		
Jeep					3	
Kia	1	8			1	
Land					3	
Lexus		6		1	3	
Lincoln		7			2	
Mazda	1	3	2	4	1	
Mercedes-Benz		16		5	2	
Mercury	1	6			1	
Mini		2				
Mitsubishi		6		3	3	1
Nissan	2	7	2	2	3	1

Break by

Brand	AWD	FWD	RWD
Acura	1	5	1
Audi	12	7	
BMW	5		15
Buick	1	8	
Cadillac	1	5	2
Chevrolet	5	18	4
Chrysler		13	2
CMC		1	
Dodge	2	7	4
Ford	4	12	7
GMC	3	1	3
Honda	3	13	1
Hummer	1		
Hyundai		12	
Infiniti		5	
Isuzu		2	
Jaguar		3	
Jeep		3	
Kia		1	
Land		3	
Lexus		4	
Lincoln		2	
Mazda		4	
Mercedes-Benz		7	
Mercury		2	
Mini		2	
Mitsubishi	2	11	
Nissan	2	12	3

Break by

Brand	3	4	5	6	8	10	12	-1
Acura		2		5				
Audi		4		10	5			
BMW				16	4			
Buick				9				
Cadillac				1	7			
Chevrolet		7		13	7			
Chrysler		5		10				
CMC					1			
Dodge		5		6	1	1		
Ford		7		7	8	1		
GMC		1		3	3			
Honda	1	11		5				
Hummer					1			
Hyundai		6		6				
Infiniti				5	3			
Isuzu				2				
Jaguar				3	9			
Jeep		1		2				
Kia		7		4				
Land				1	2			
Lexus				6	5			
Lincoln				2	7			
Mazda		7		2				2
Mercedes-Benz		2		12	10		2	
Mercury				5	4			
Mini		2						
Mitsubishi		7		6				
Nissan		4		11	2			

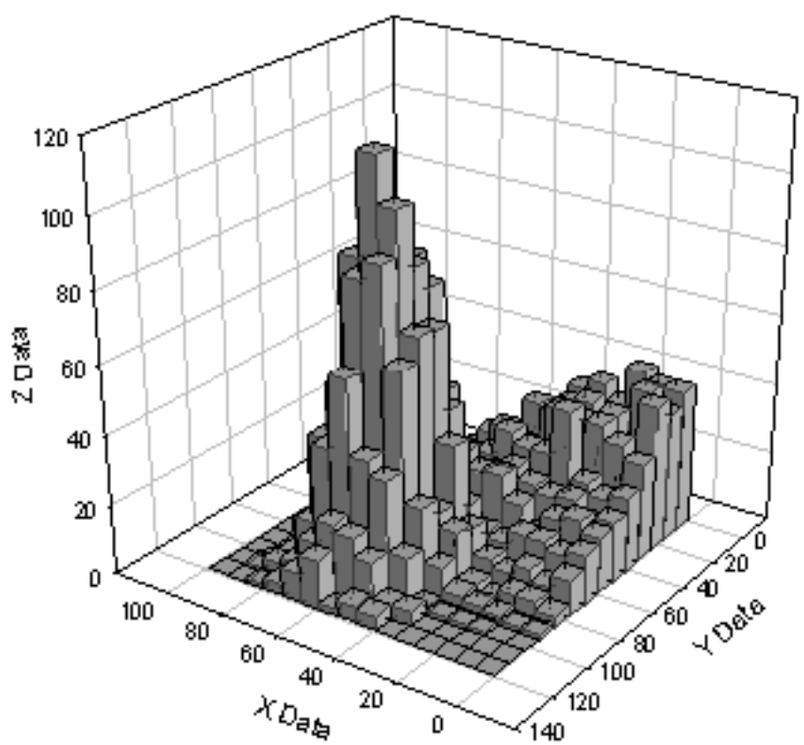
Create a sheet with a parameter that lets you select what is the other dimension, chosen from:

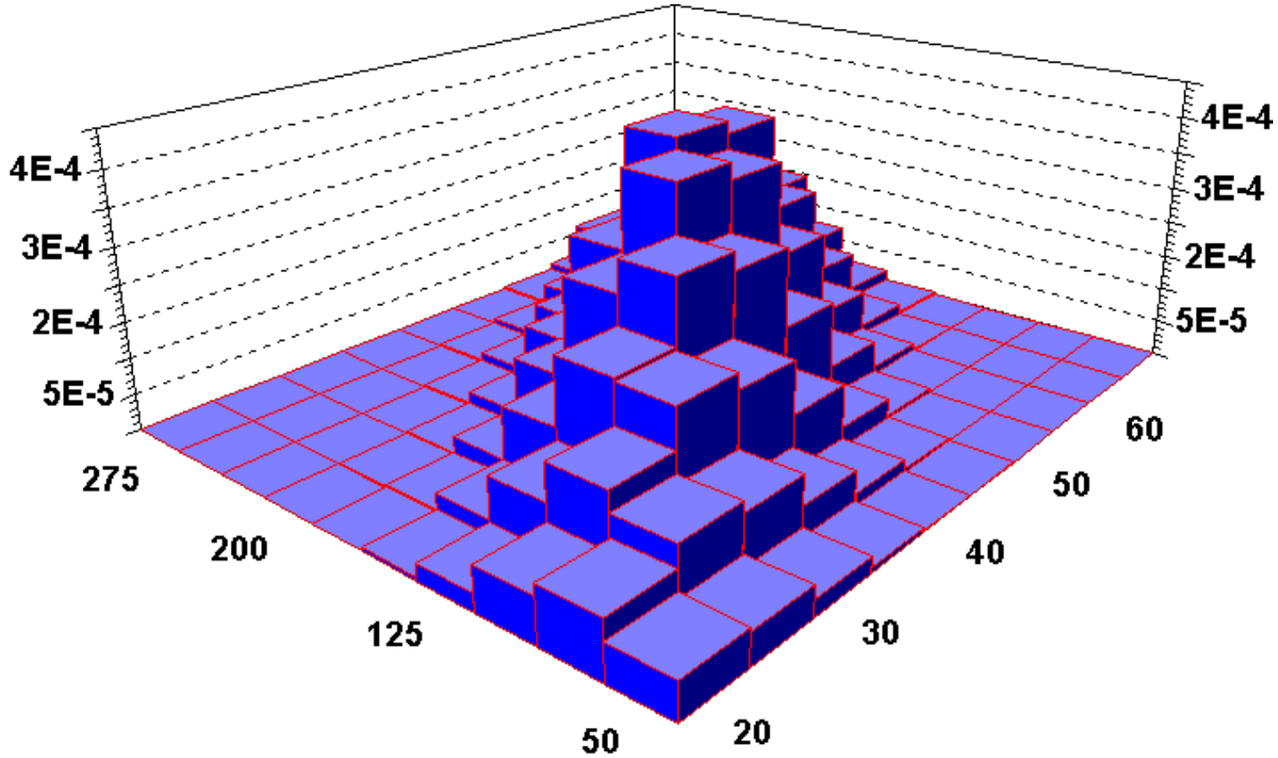
1. Class
2. Drive Train
3. Cyl

Joint Distribution of (2 categorical) variables

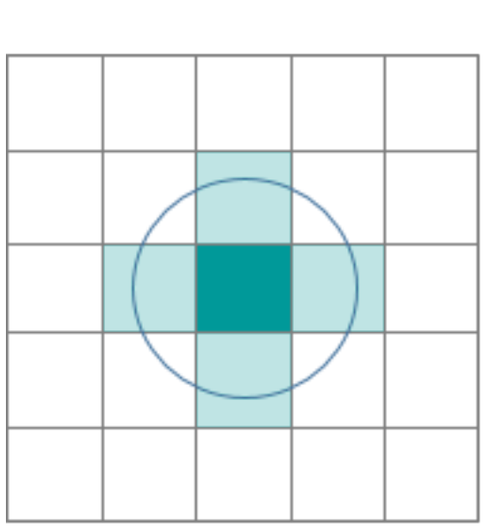
- **Create a Parameter that will take a string value from a list of variable names to be used to break your distribution**
  - **ex: Class, Engine Type and Cyl**
- **Create a Calculated Field that will take the value of one of your initial variables (Class, Engine Type and Cyl) depending on the value of the parameter**
- **Use your new Calculated Field in the remaining empty shelf (the other is already occupied by the other dimension - Brand)**

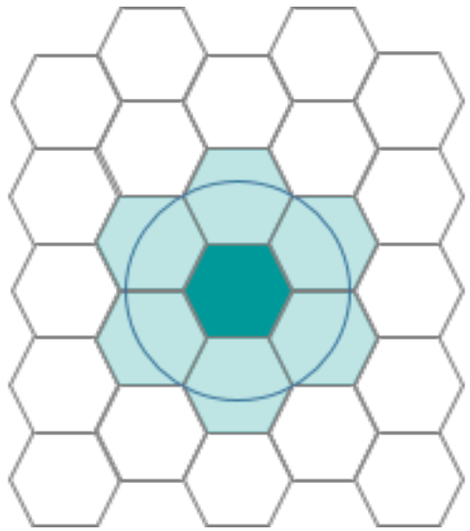
Joint Distribution of (2 continuous) variables





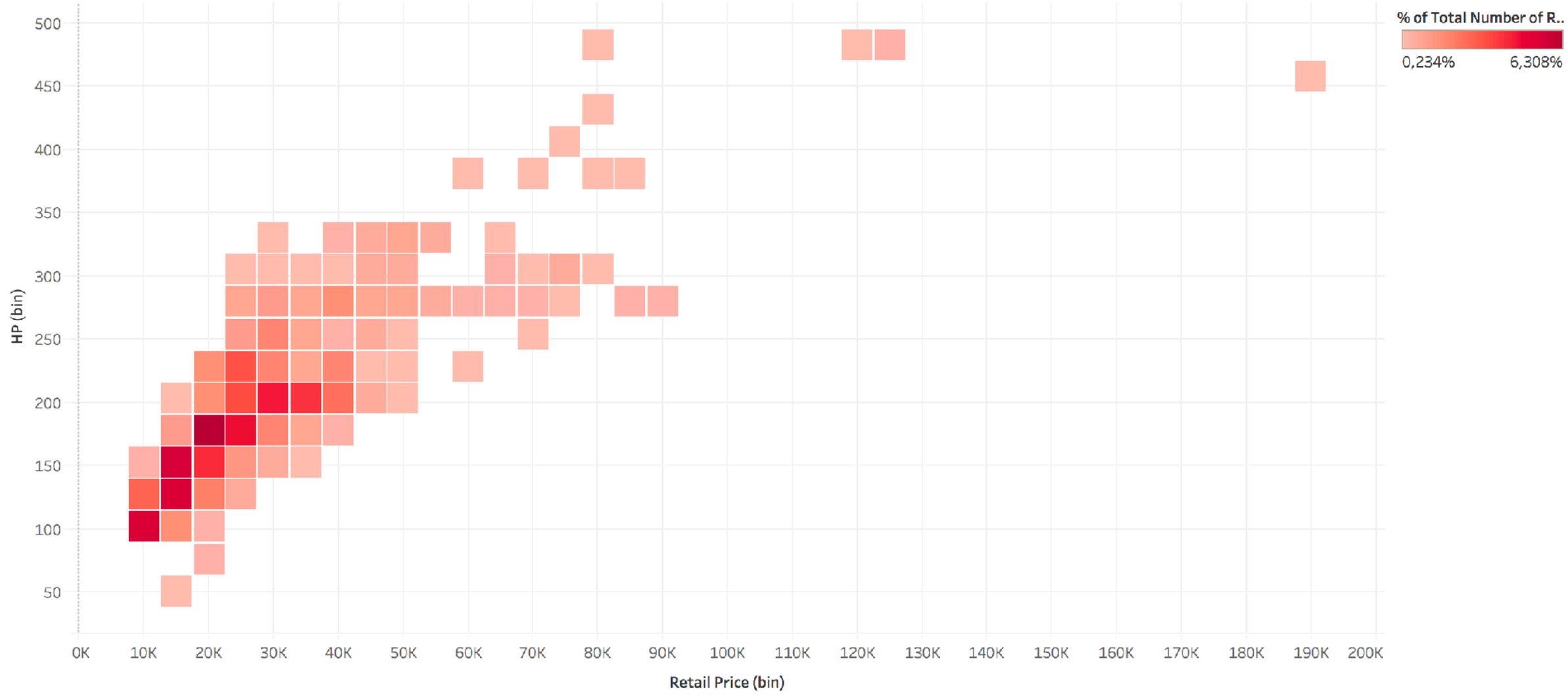
Joint Distribution of (2 continuous) variables





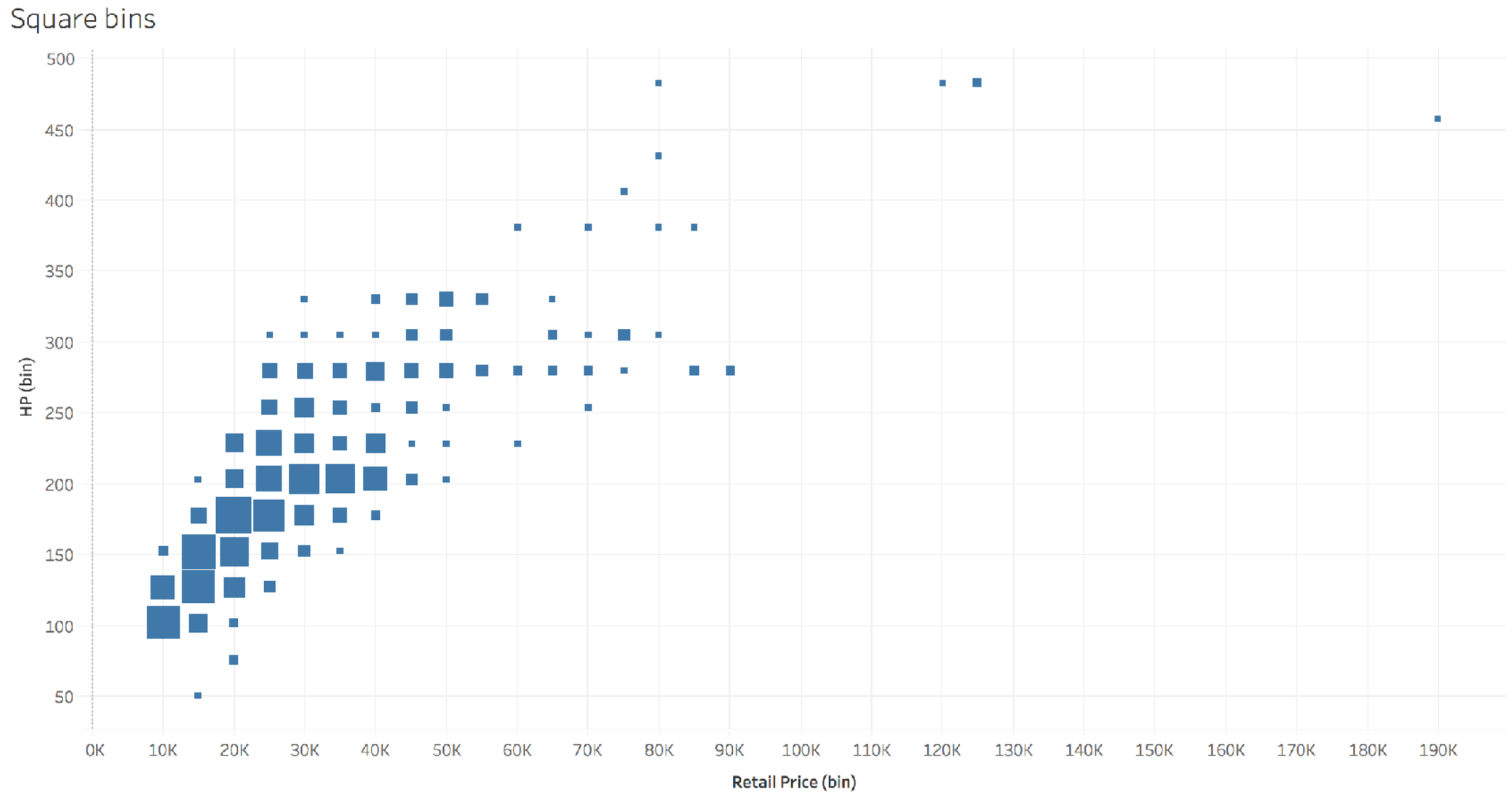
Joint Distribution of (2 continuous) variables

Square bins



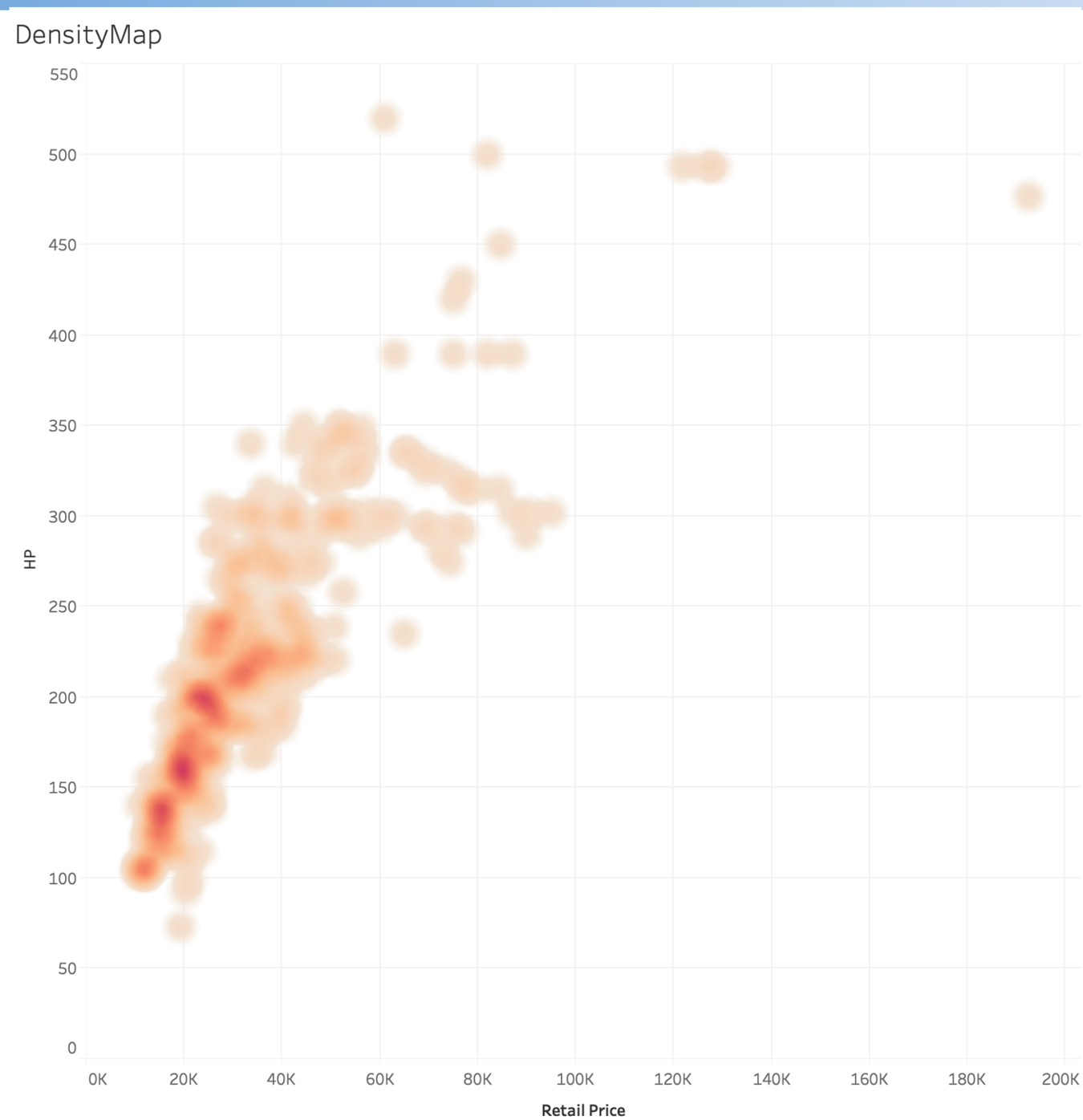
Retail Price (bin) vs. HP (bin). Color shows % of Total Number of Records.

Joint Distribution of (2 continuous) variables



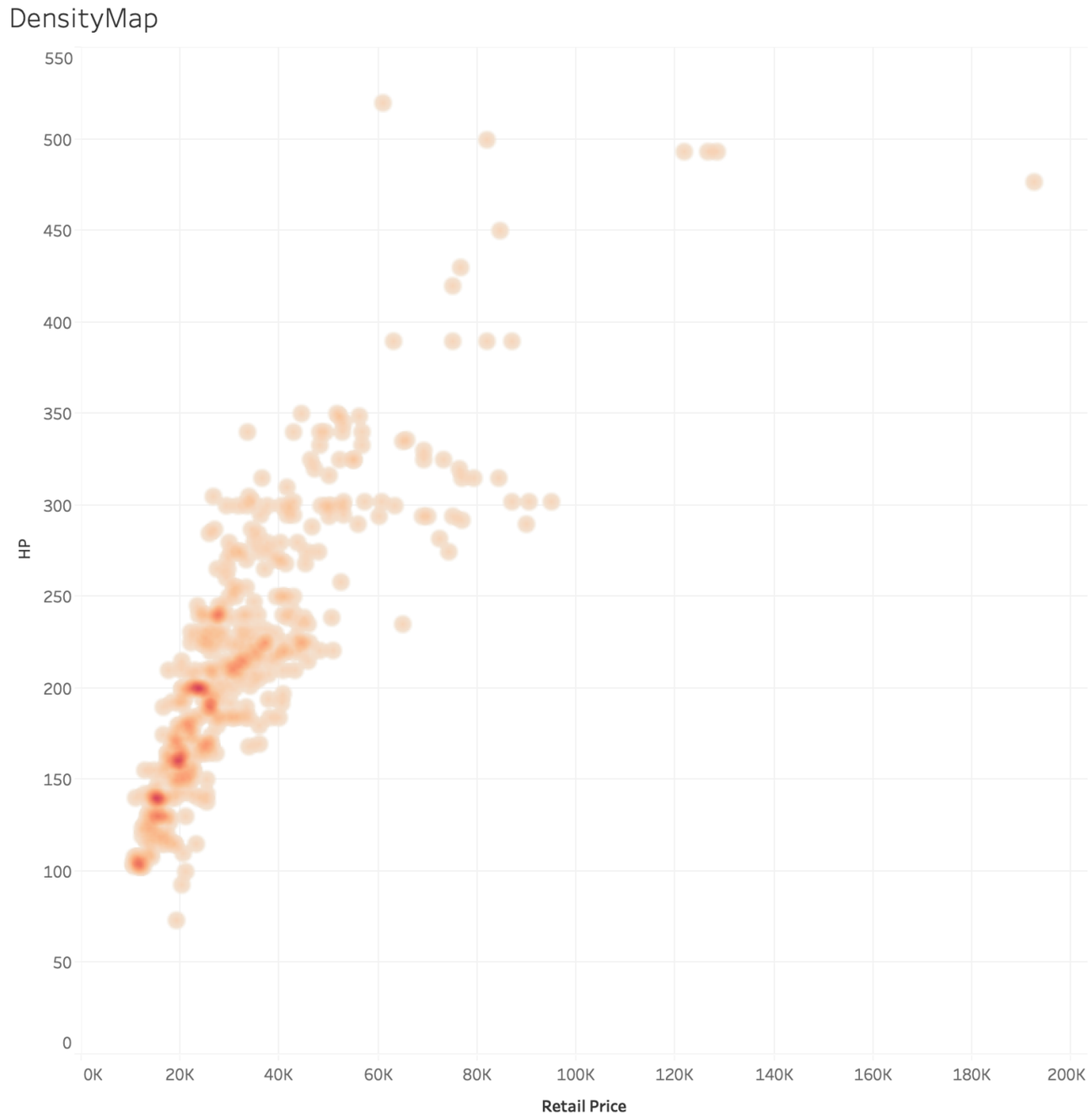
Retail Price (bin) vs. HP (bin). Size shows % of Total Number of Records.

Joint Distribution of (2 continuous) variables



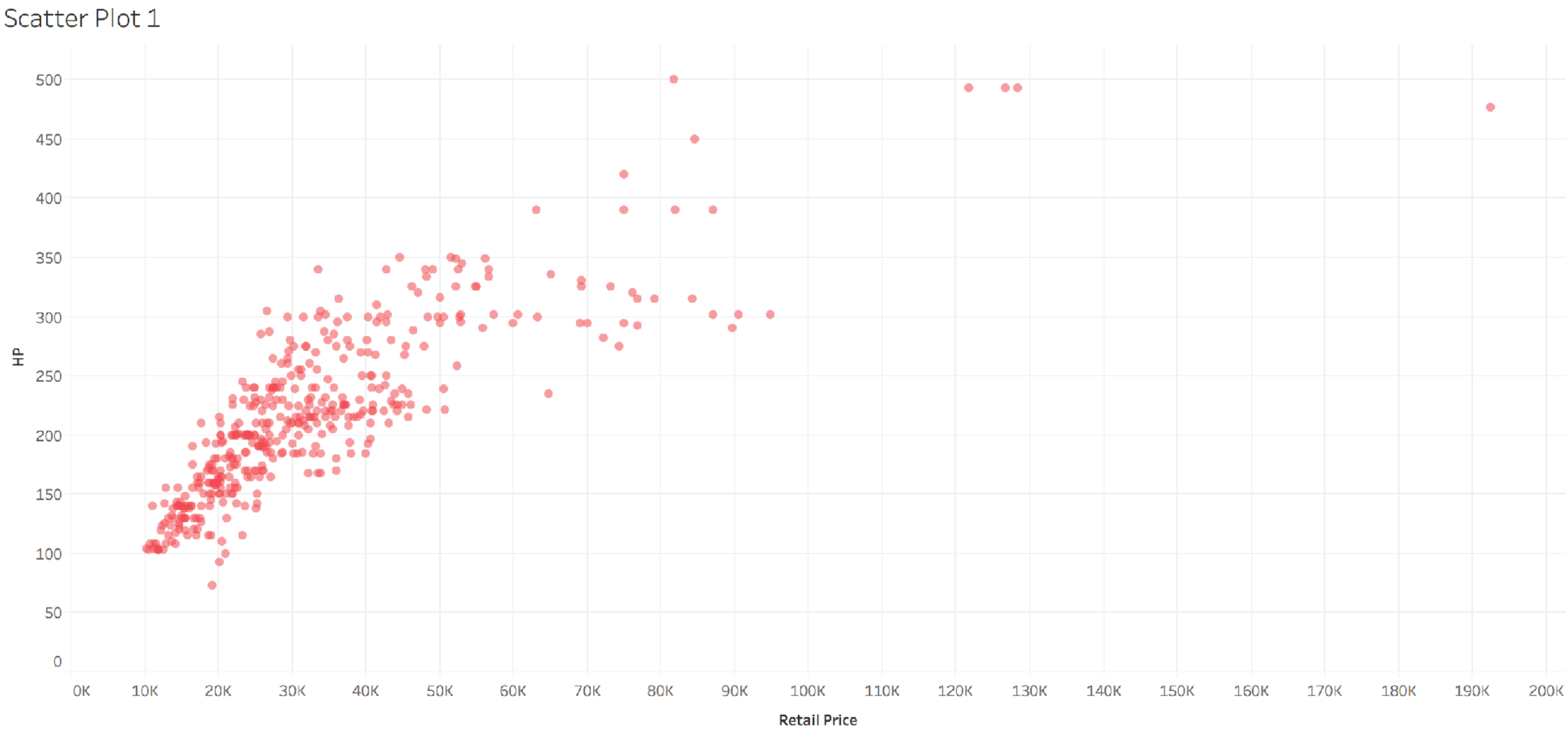
Sum of Retail Price vs. sum of HP. Details are shown for Vehicle Name.

Joint Distribution of (2 continuous) variables



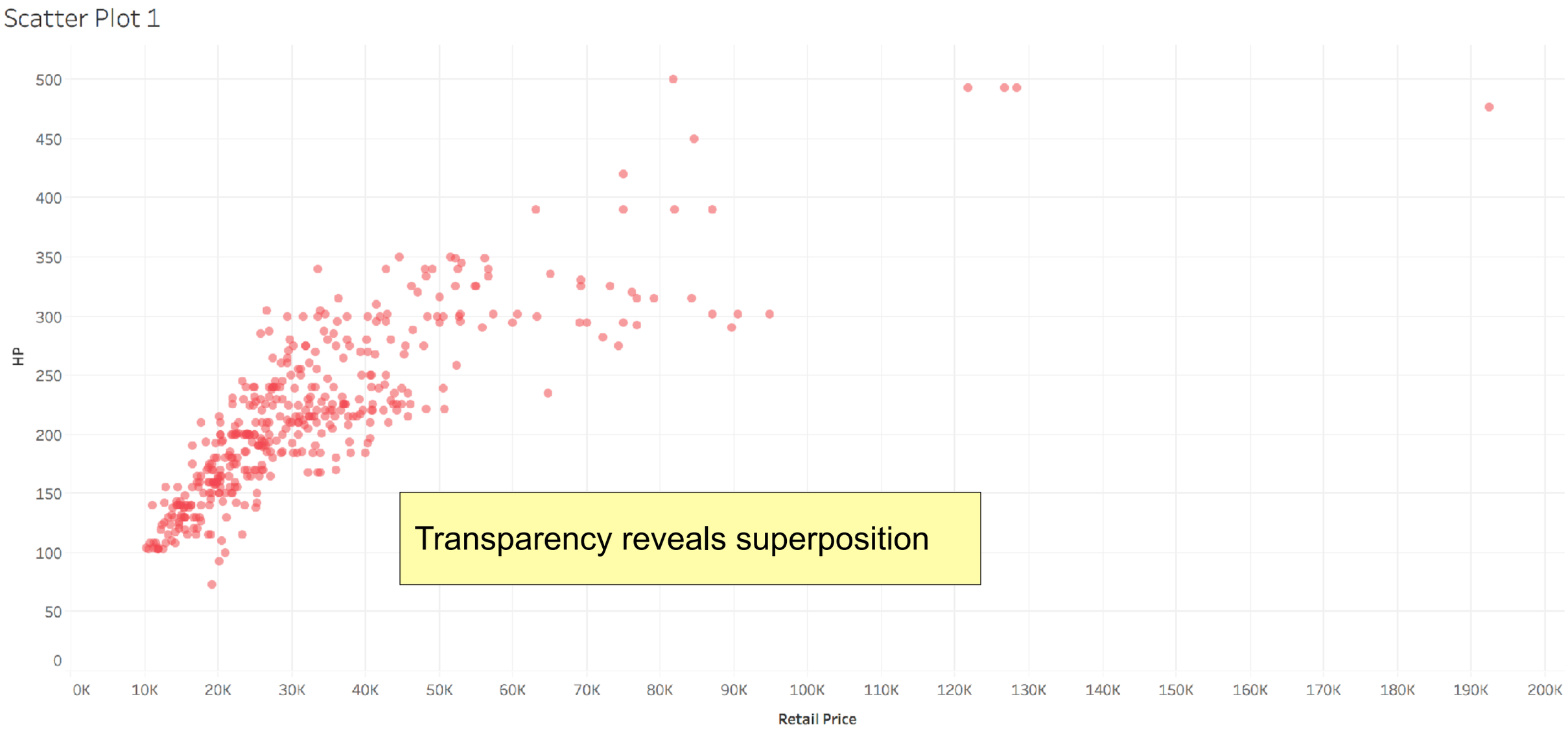
Sum of Retail Price vs. sum of HP. Details are shown for Vehicle Name.

Joint Distribution of (2 continuous) variables



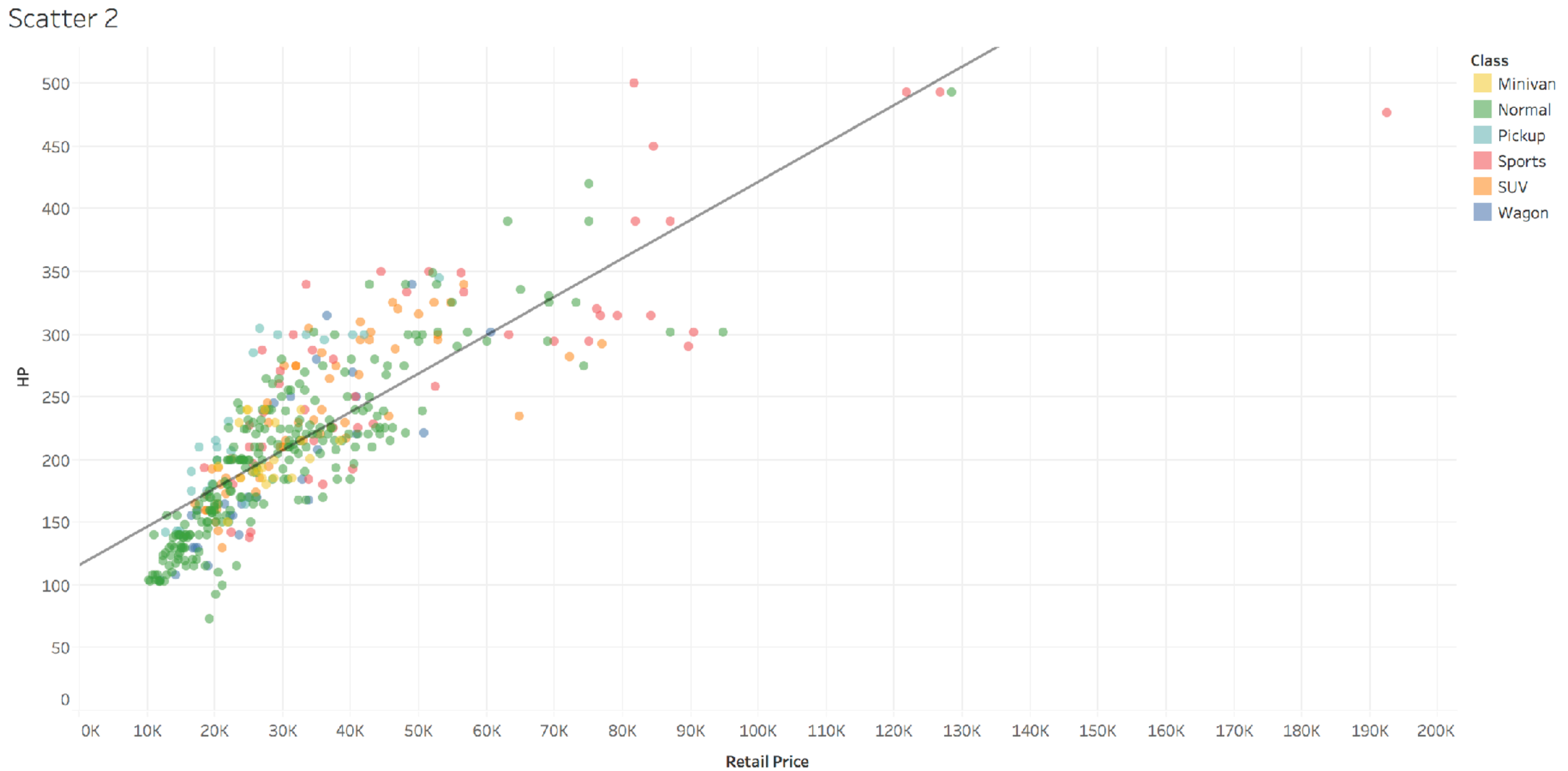
Retail Price vs. HP.

Joint Distribution of (2 continuous) variables



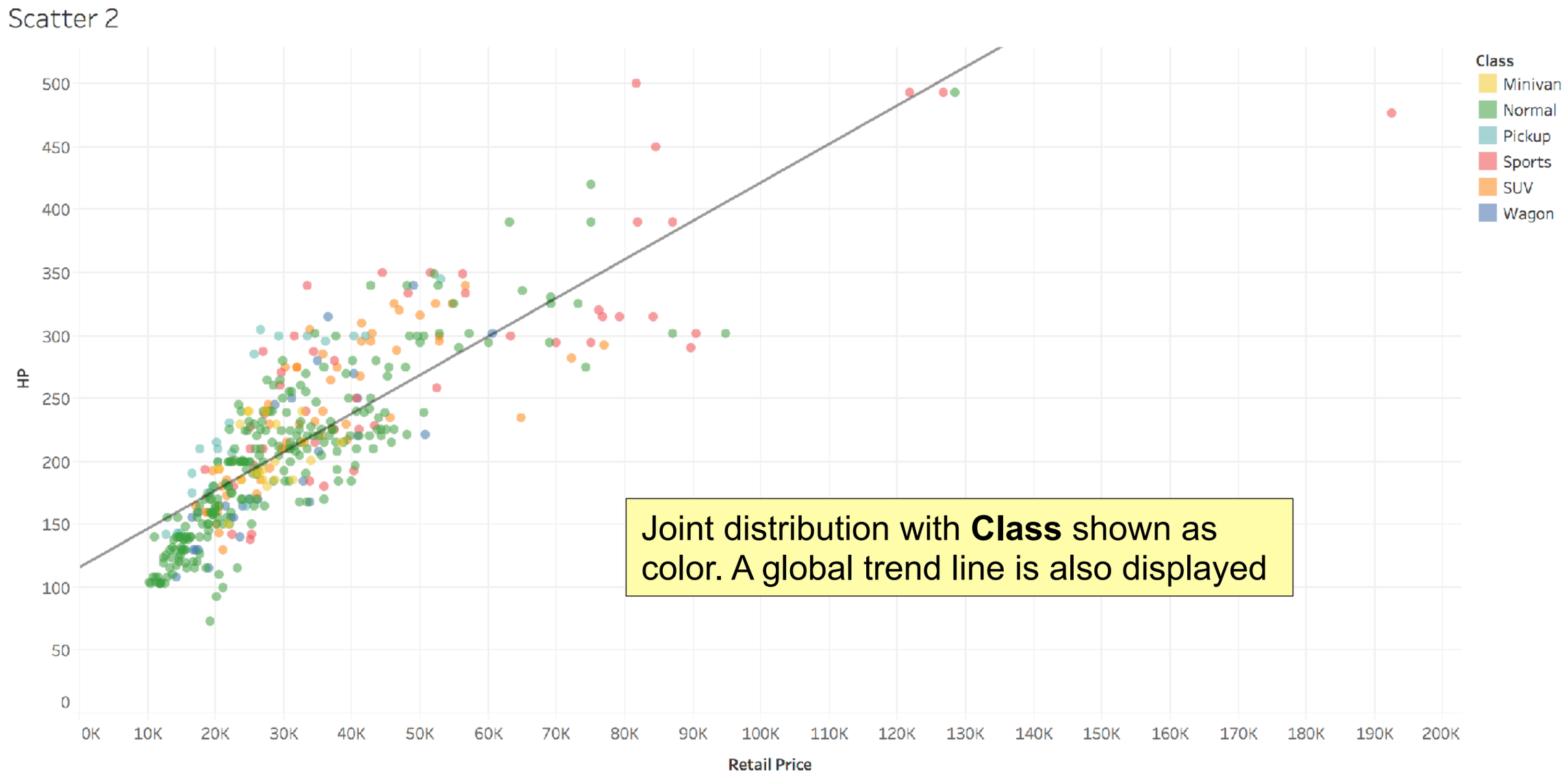
Retail Price vs. HP.

Joint Distribution of (2 continuous) variables



Retail Price vs. HP. Color shows details about Class.

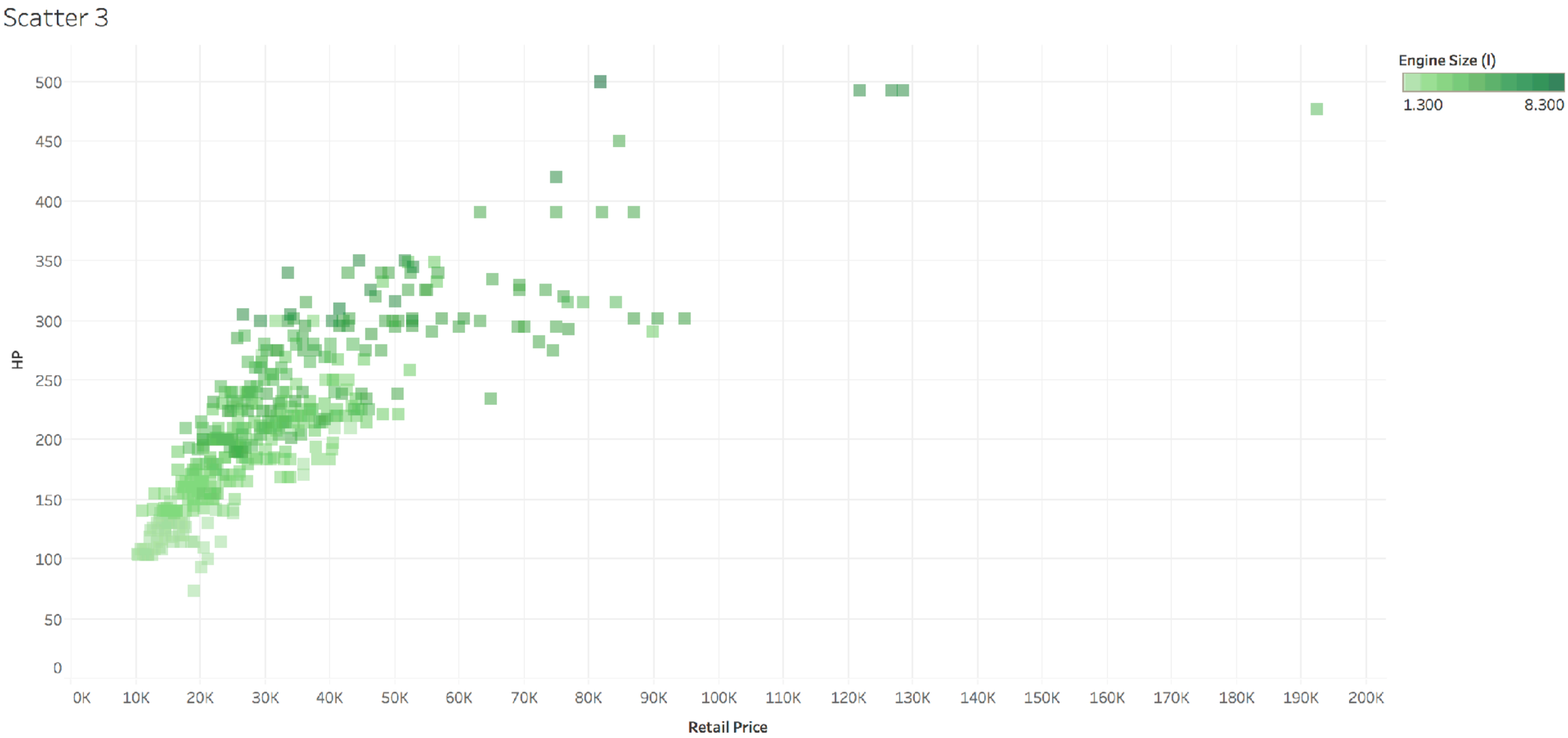
Joint Distribution of (2 continuous) variables



Joint distribution with **Class** shown as color. A global trend line is also displayed

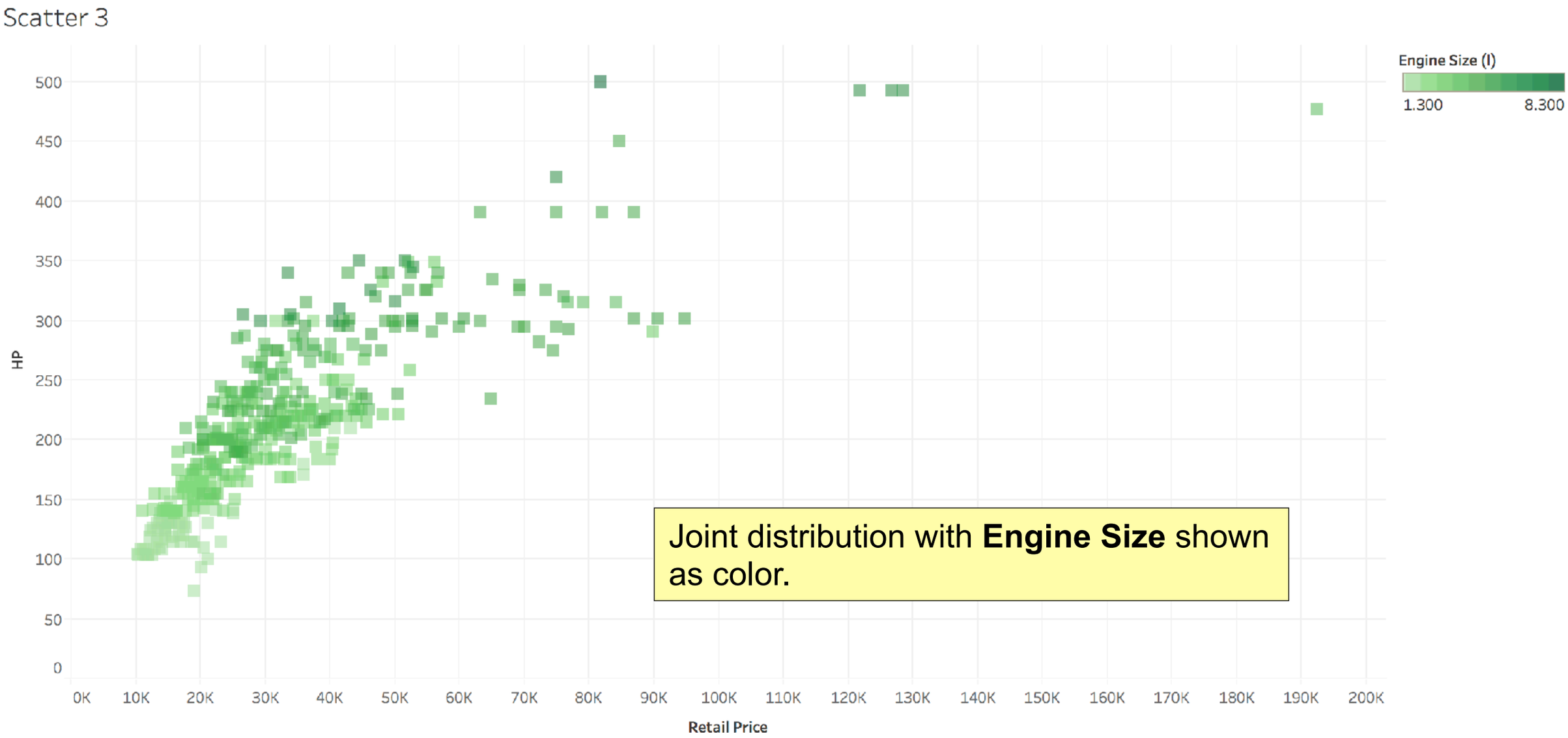
Retail Price vs. HP. Color shows details about Class.

Joint Distribution of (2 continuous) variables



Retail Price vs. HP. Color shows Engine Size (l).

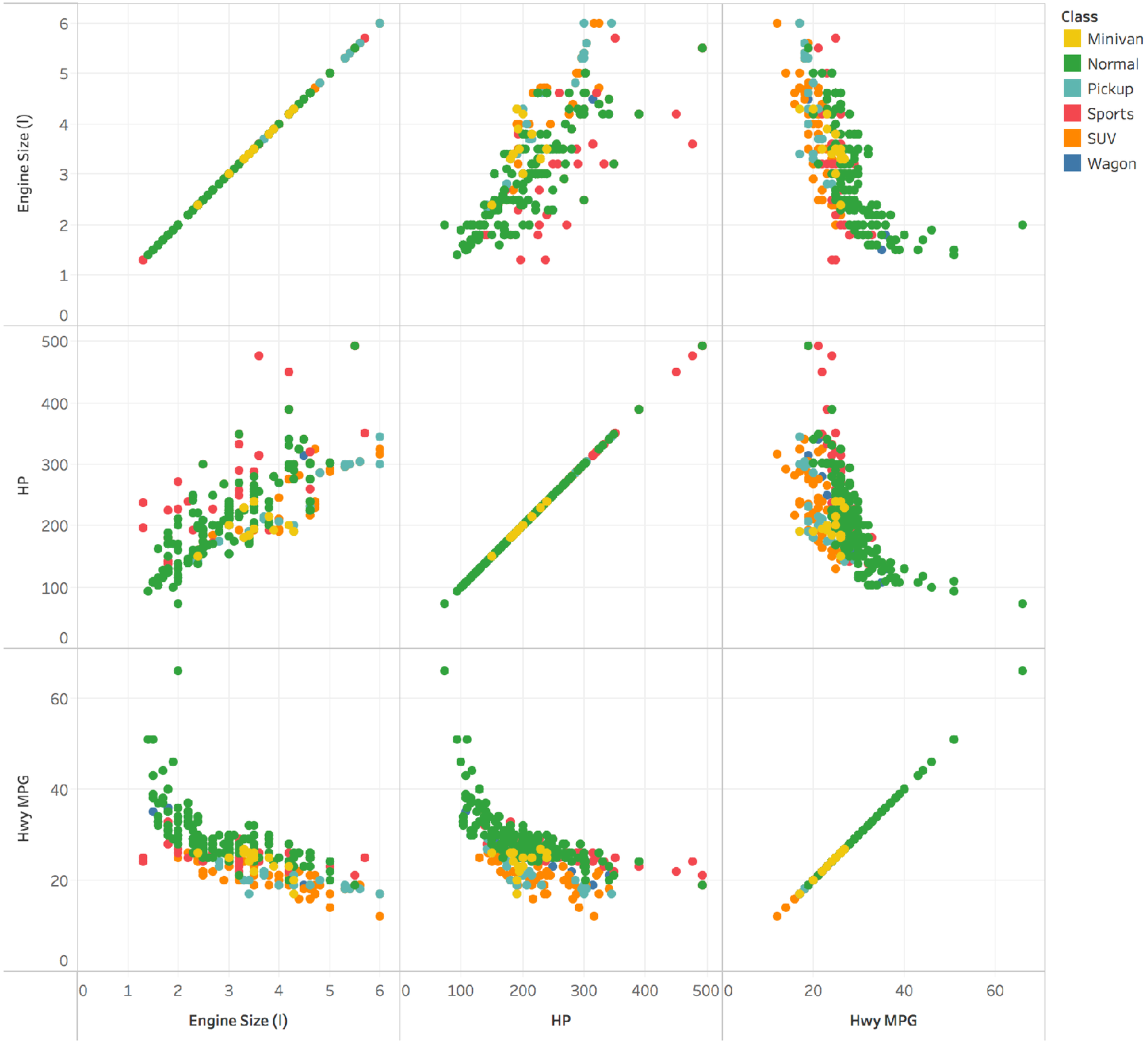
Joint Distribution of (2 continuous) variables



Retail Price vs. HP. Color shows Engine Size (l).

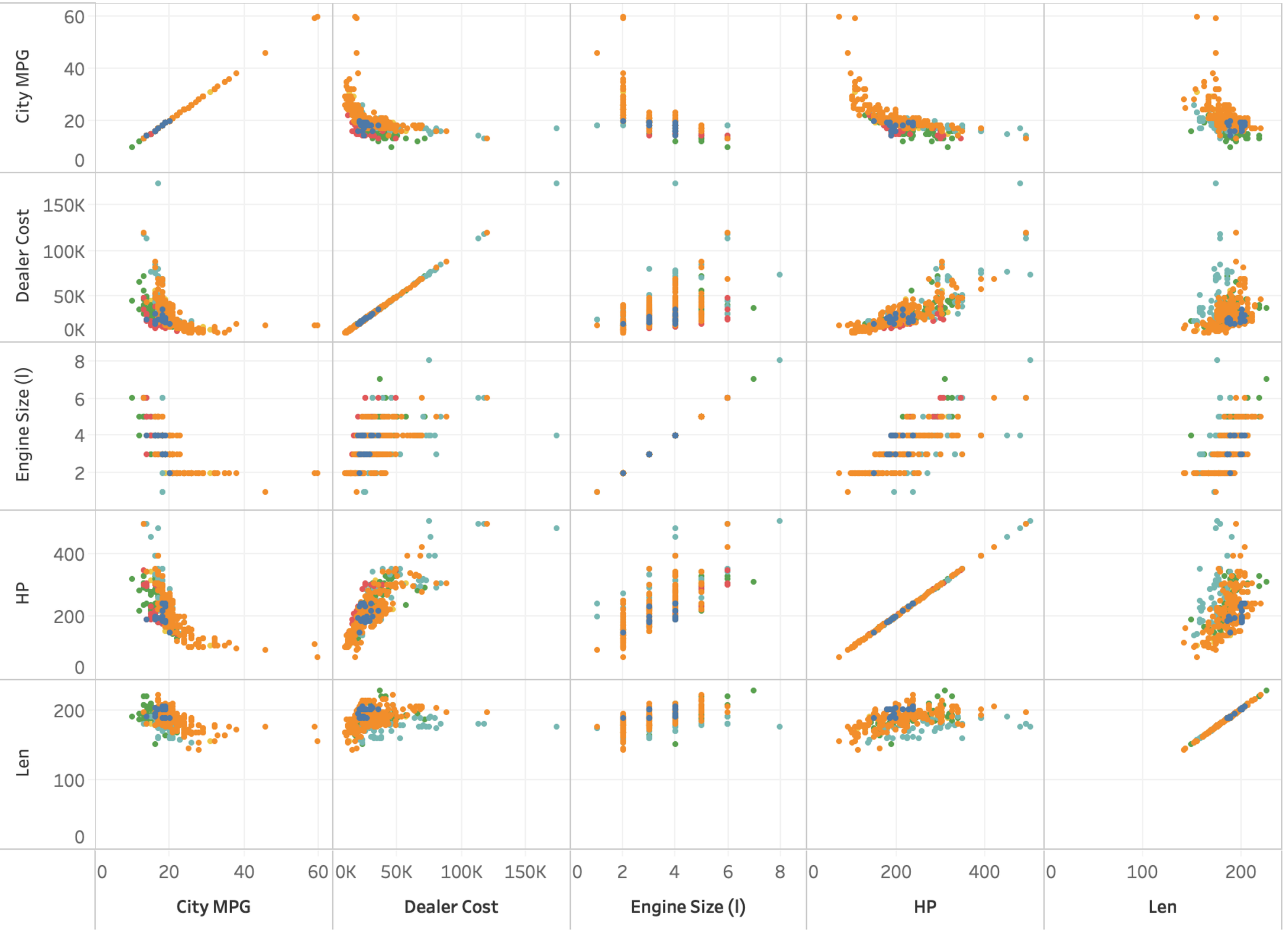
Scatter Plot Matrix

Scatter Matrix



Engine Size (l), HP and Hwy MPG vs. Engine Size (l), HP and Hwy MPG. Color shows details about Class. The data is filtered on Hwy MPG, which keeps non-Null values only.

Matrix of Scatter Matrix



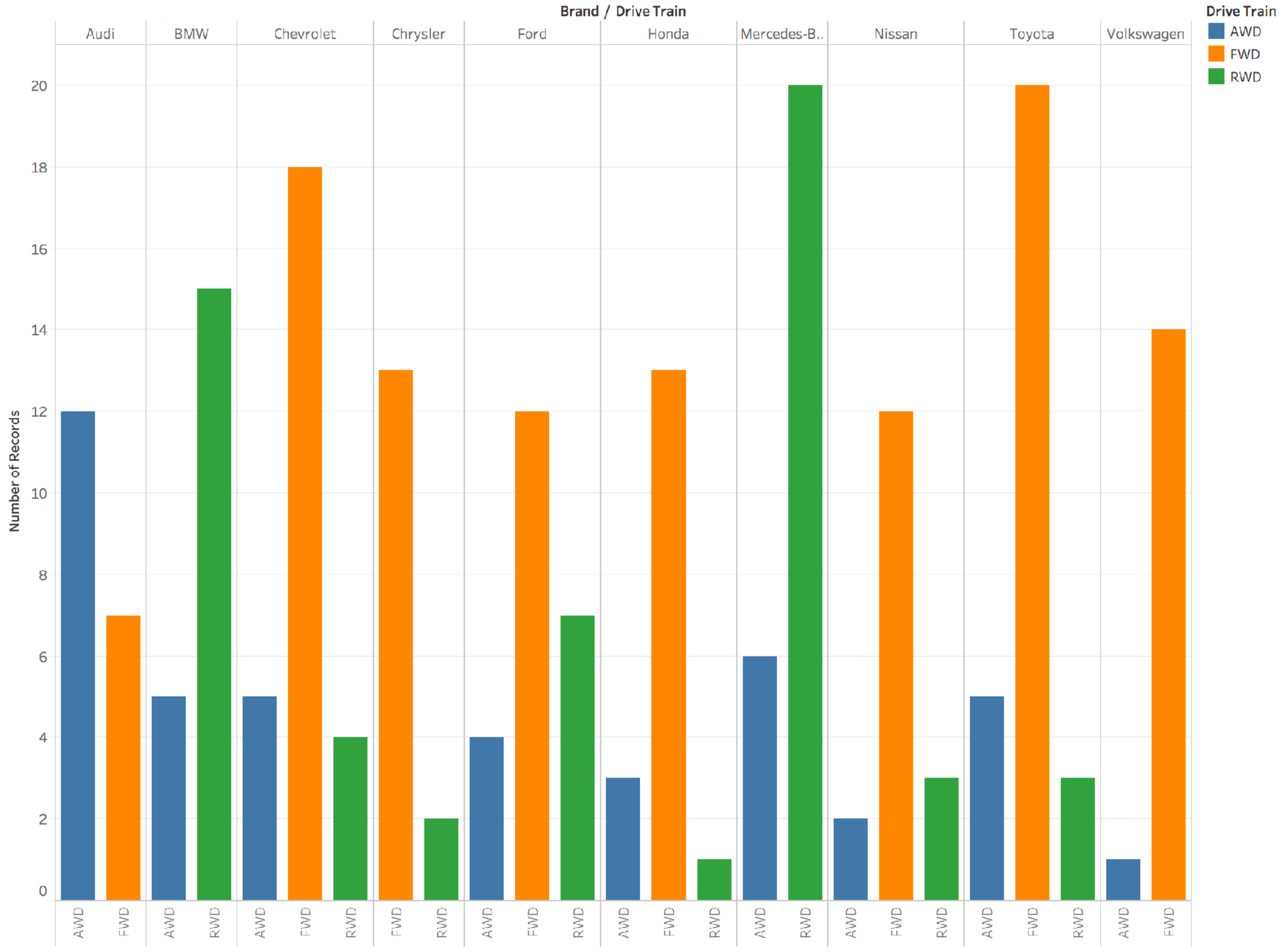
City MPG, Dealer Cost, Engine Size (l), HP and Len vs. City MPG, Dealer Cost, Engine Size (l), HP and Len. Color shows details about Class. Details are shown for Vehicle Name.

Explore the Data: One Variable Split by a Categorical

Hands On: Cars Dataset

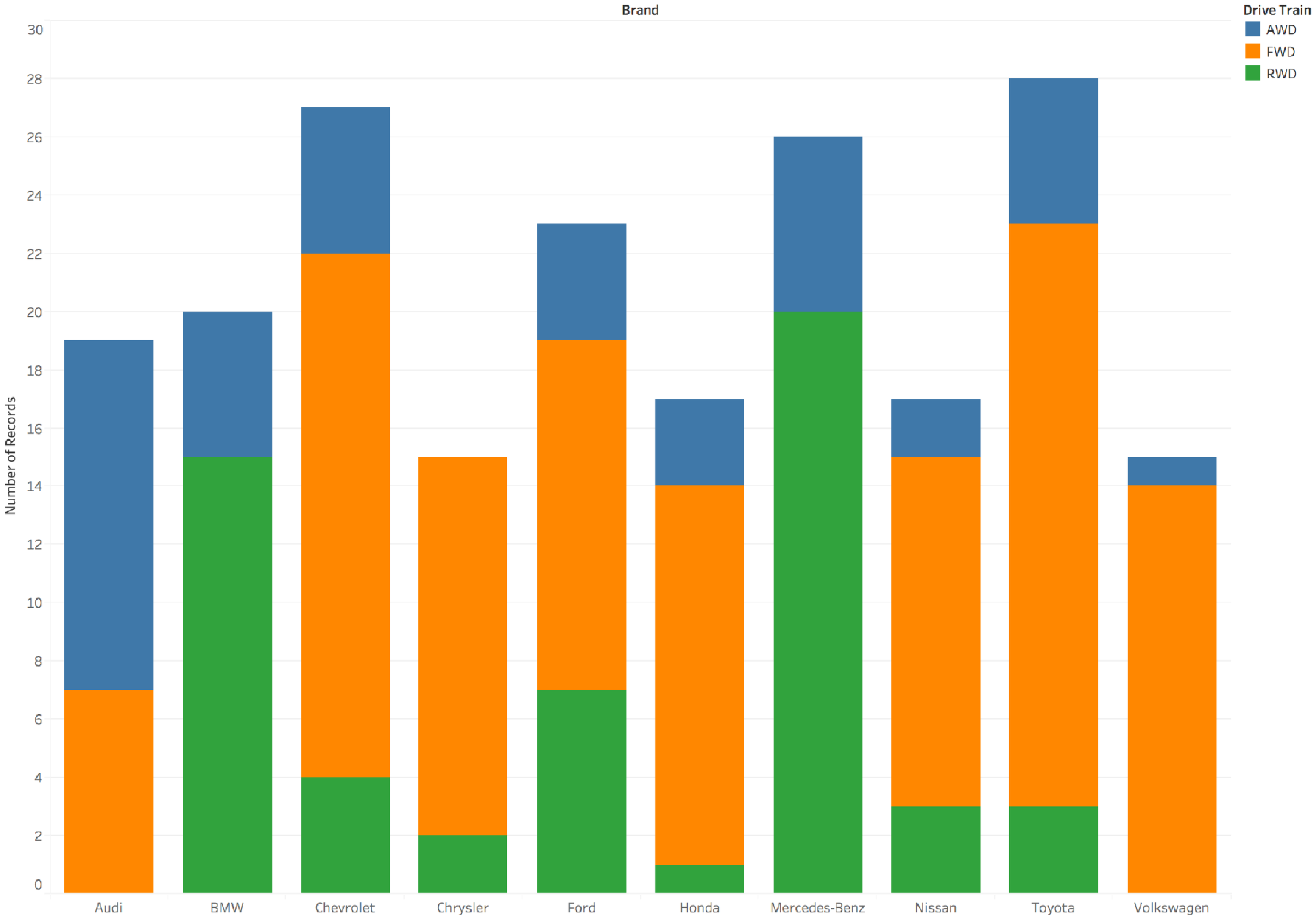
- **Sometimes we want to see our distribution broken down by additional categorical variables:**
  - **Sales by product type**
  - **Sales by quarter**
  - **Models by Engine Type**
  - **...**

Absolute Side by side



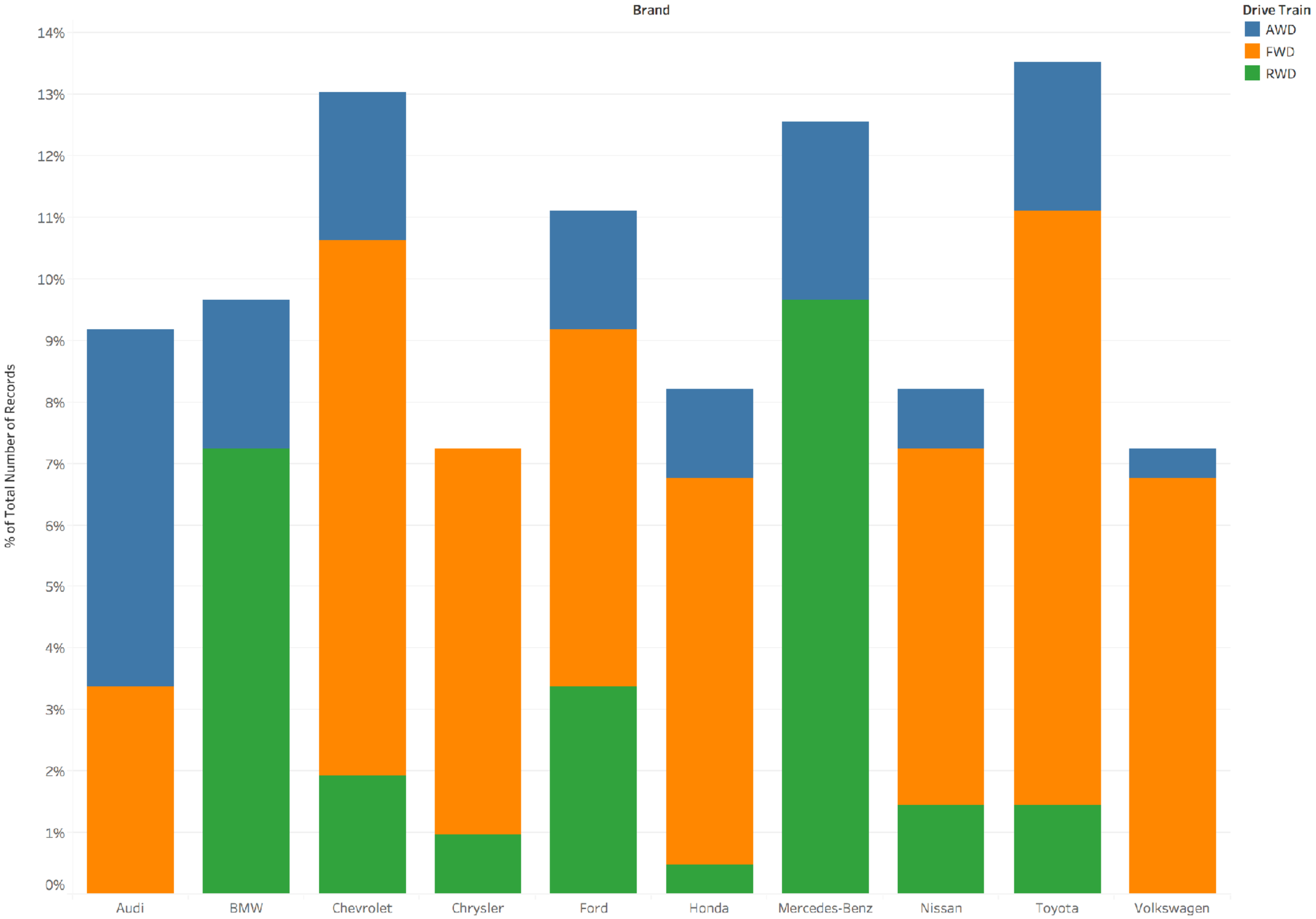
Sum of Number of Records for each Drive Train broken down by Brand. Color shows details about Drive Train. The view is filtered on Brand, which keeps 10 of 39 members.

Absolute Stacked



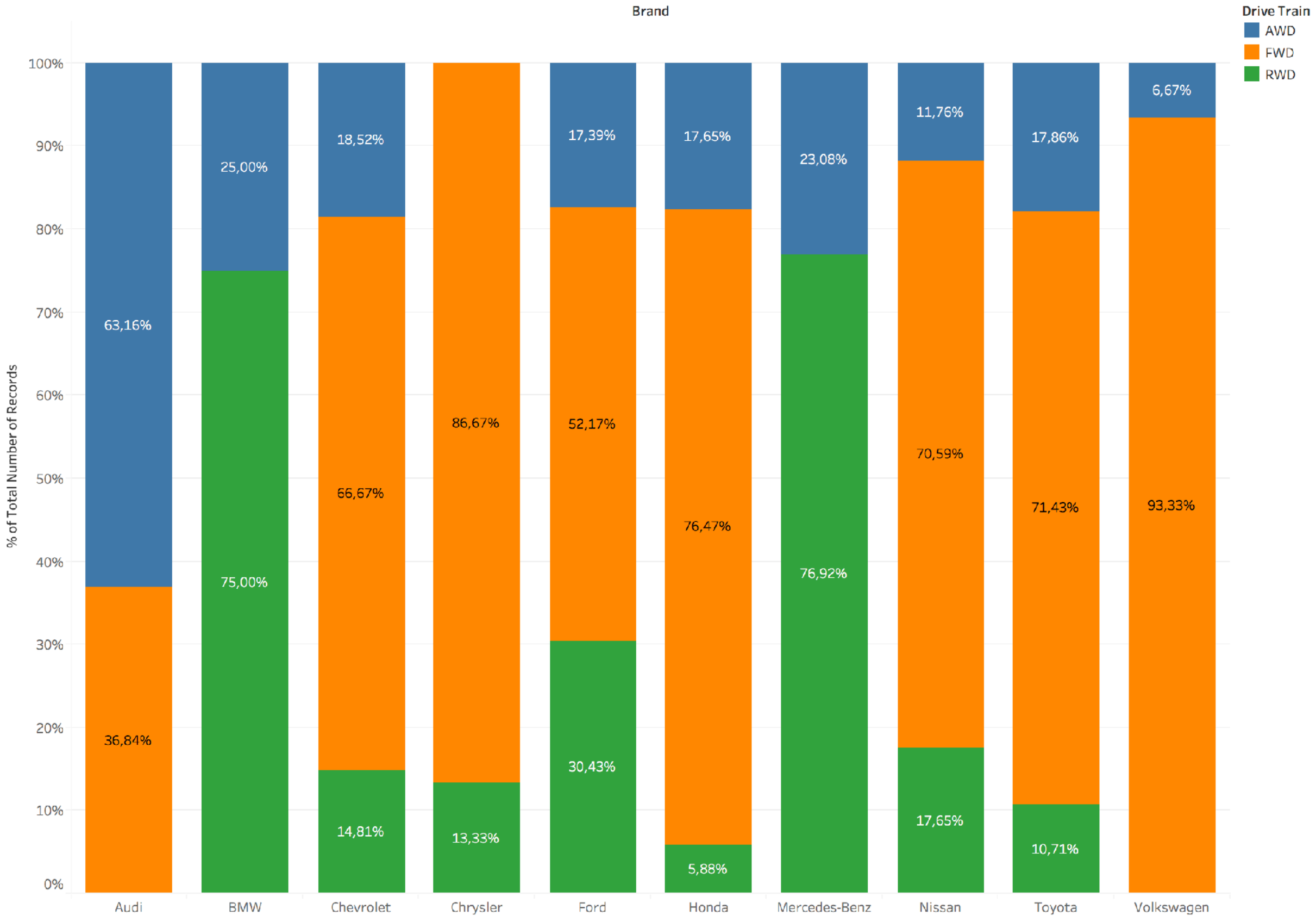
Sum of Number of Records for each Brand. Color shows details about Drive Train. The view is filtered on Brand, which keeps 10 of 39 members.

Relative Split by



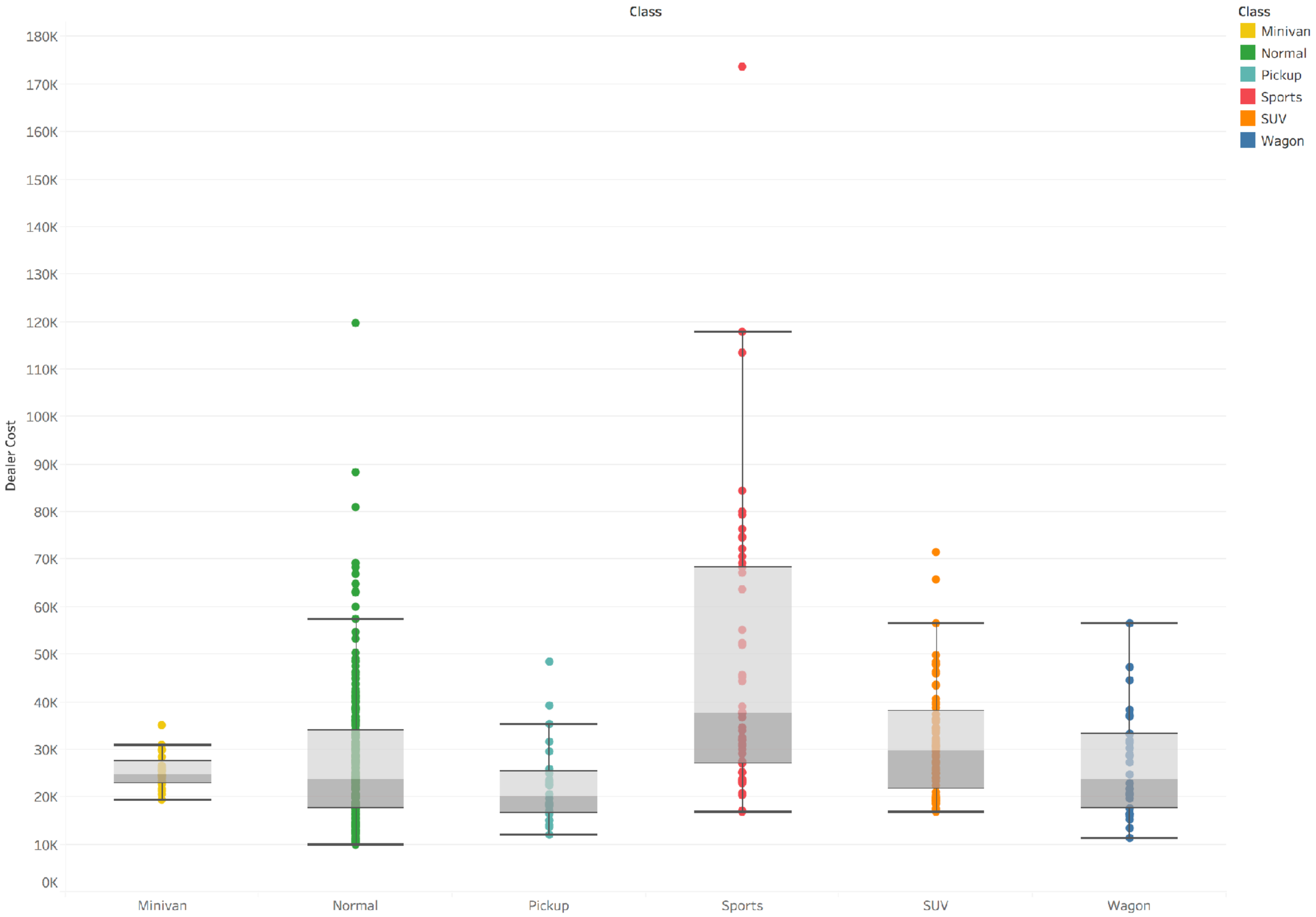
% of Total Number of Records for each Brand. Color shows details about Drive Train. The view is filtered on Brand, which keeps 10 of 39 members.

Relative 100%



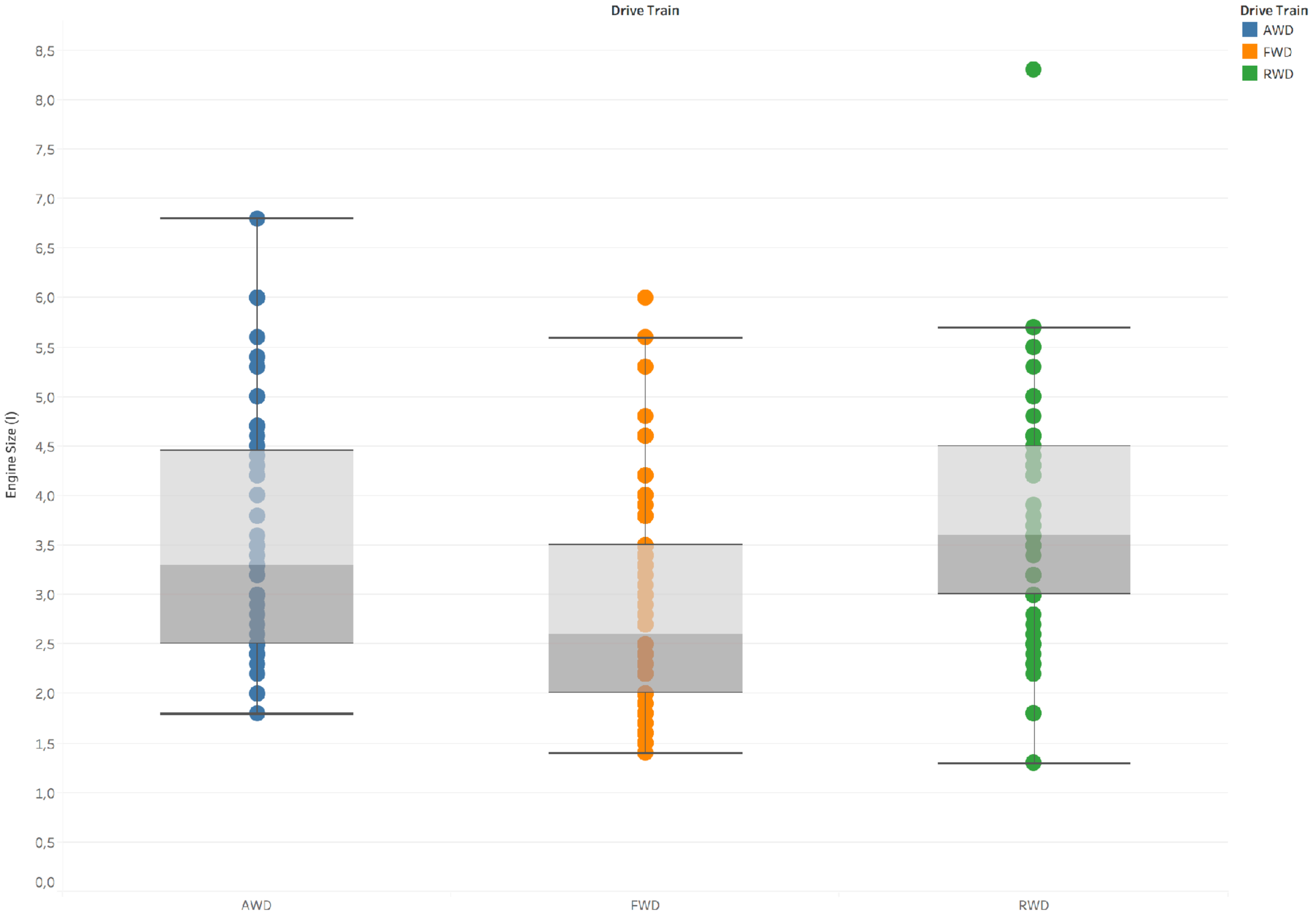
% of Total Number of Records for each Brand. Color shows details about Drive Train. The view is filtered on Brand, which keeps 10 of 39 members.

Boxplot of Dealer Cost by Class



Dealer Cost for each Class. Color shows details about Class.

Engine Size by Drive Train

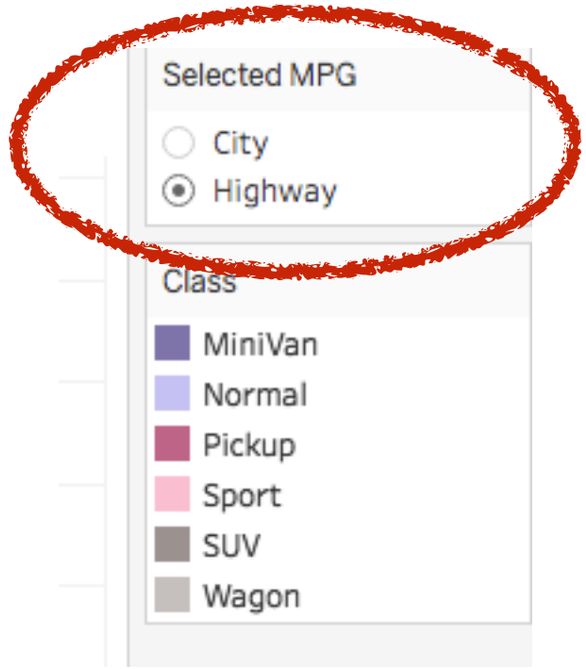


Engine Size (l) for each Drive Train. Color shows details about Drive Train. Details are shown for Drive Train.

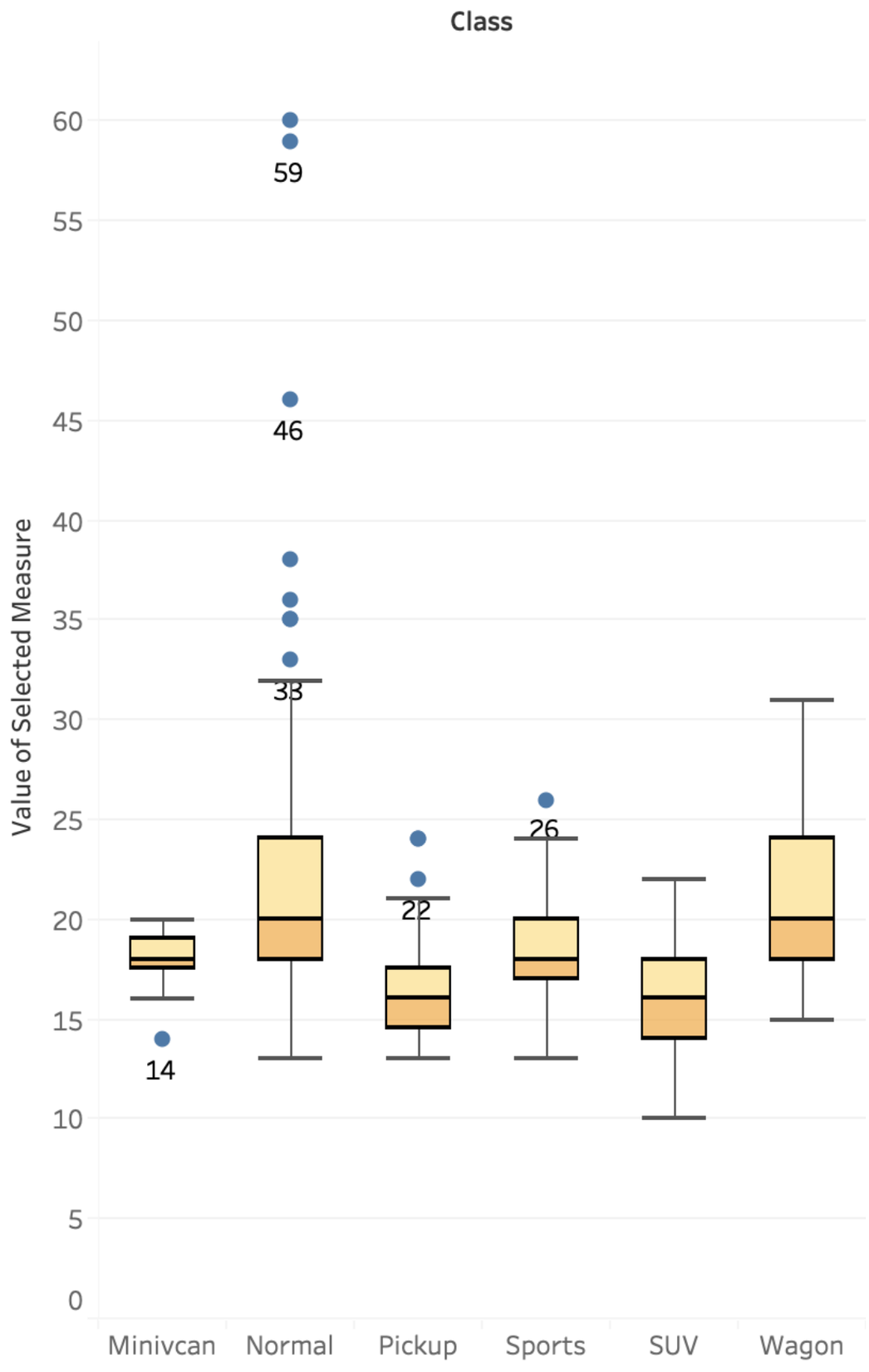
Filters and Parameters

Hands On: Cars Dataset

- **Steps:**
  - **Add interactivity to some of your visualizations using:**
    - Filters
    - Parameters
  - **Change between Retail Price vs MPG per class visualization to let the user choose between City MPG or Highway MPG.**



Box Plot With Parameter . The Selected mesasure: **City MPG**

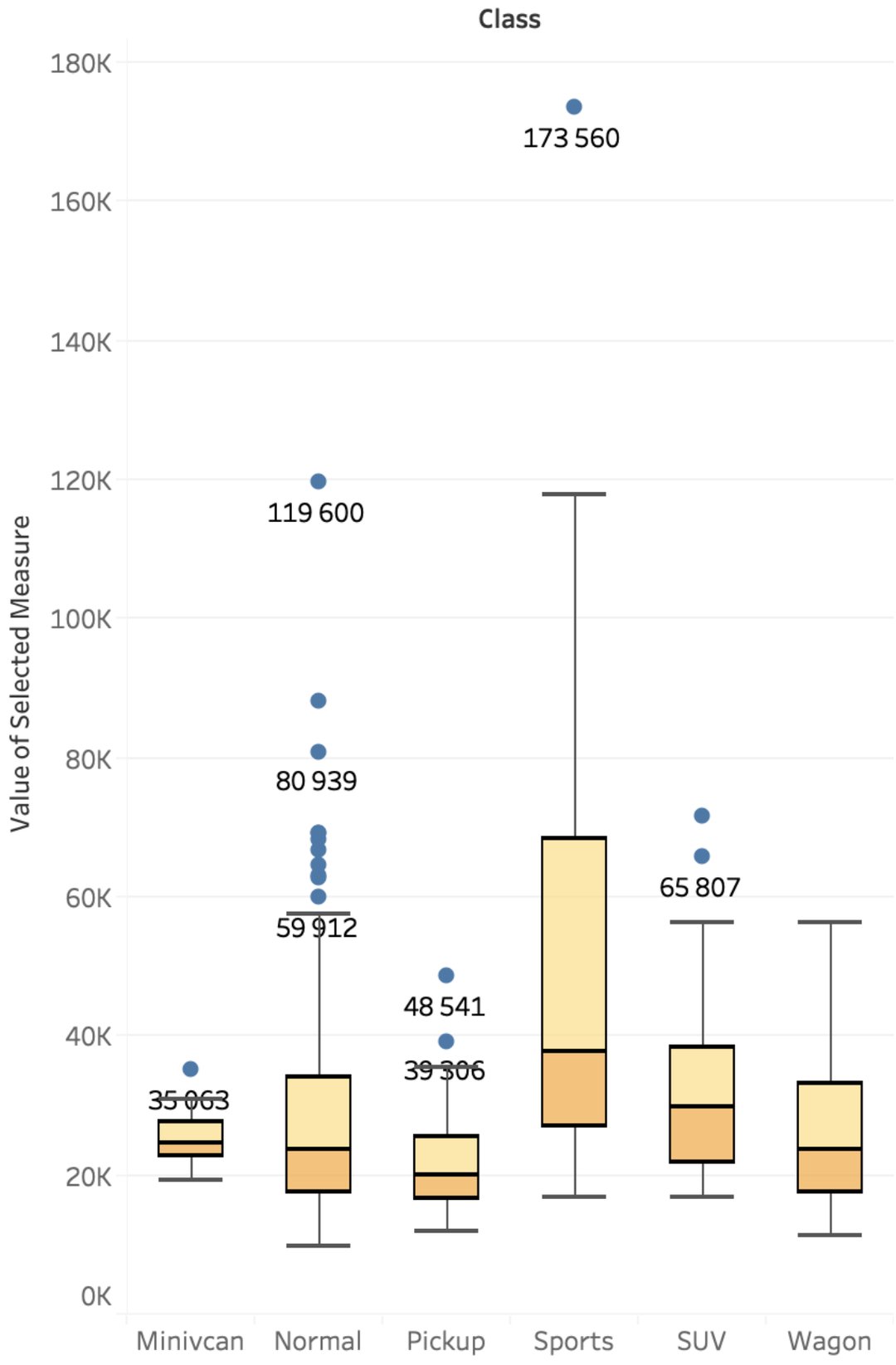


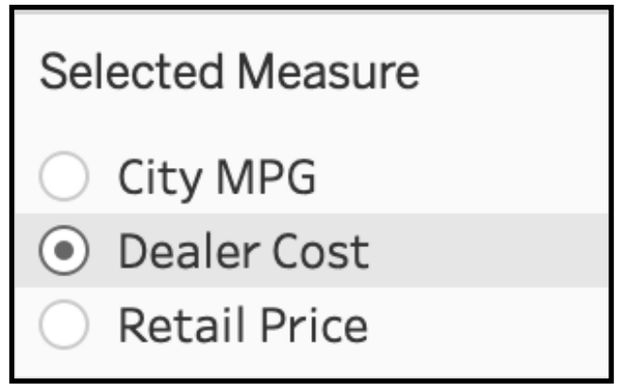
Value of Selected Measure for each Class. Details are shown for Vehicle Name.

Selected Measure

- (•) City MPG
- ( ) Dealer Cost
- ( ) Retail Price

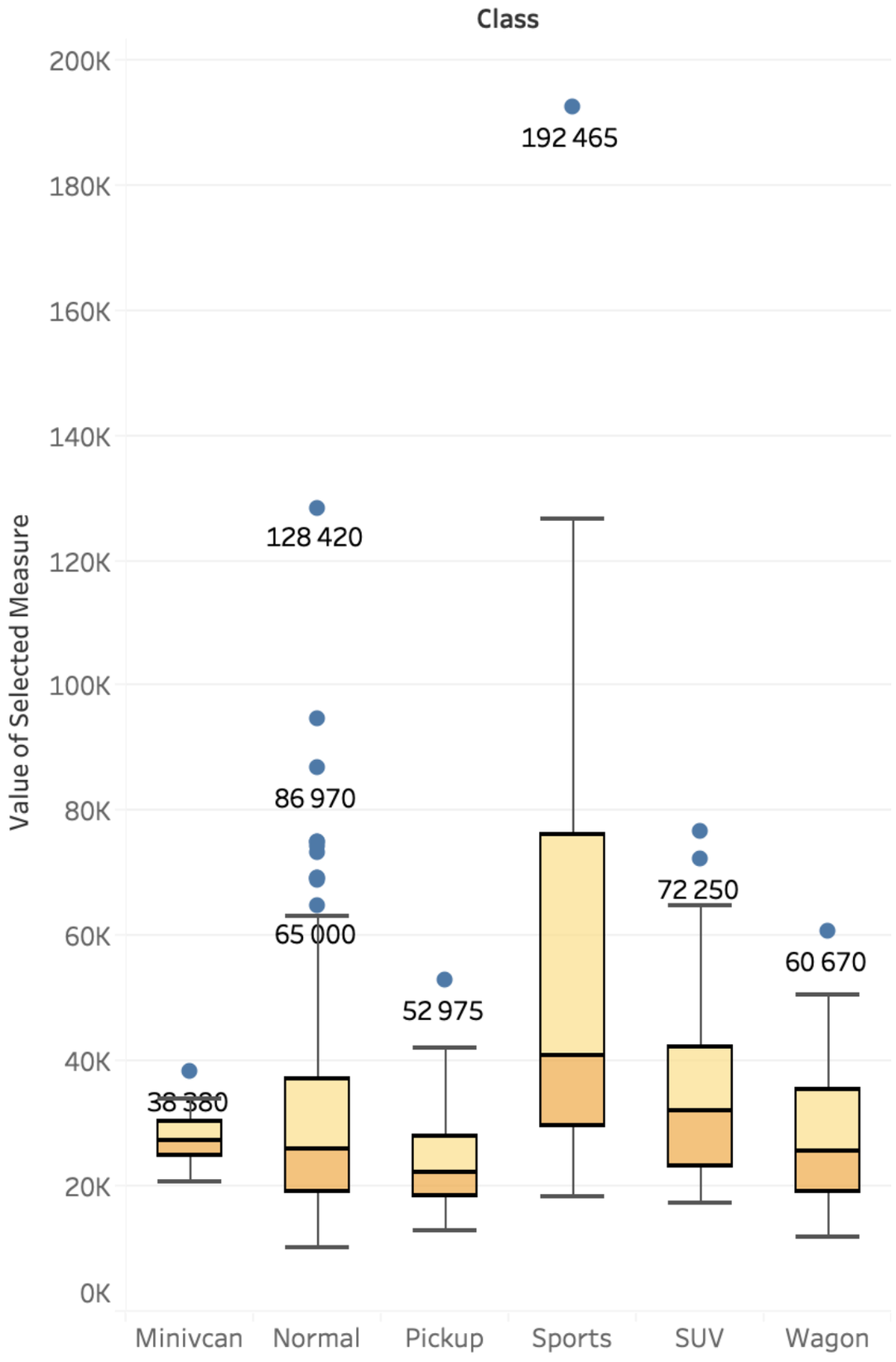
Box Plot With Parameter . The Selected mesasure: **Dealer Cost**





Value of Selected Measure for each Class. Details are shown for Vehicle Name.

Box Plot With Parameter . The Selected mesasure: **Retail Price**



Value of Selected Measure for each Class. Details are shown for Vehicle Name.

Selected Measure

- ○ City MPG
- ○ Dealer Cost
- ◉ Retail Price

Pages

Filters

Marks

Circle

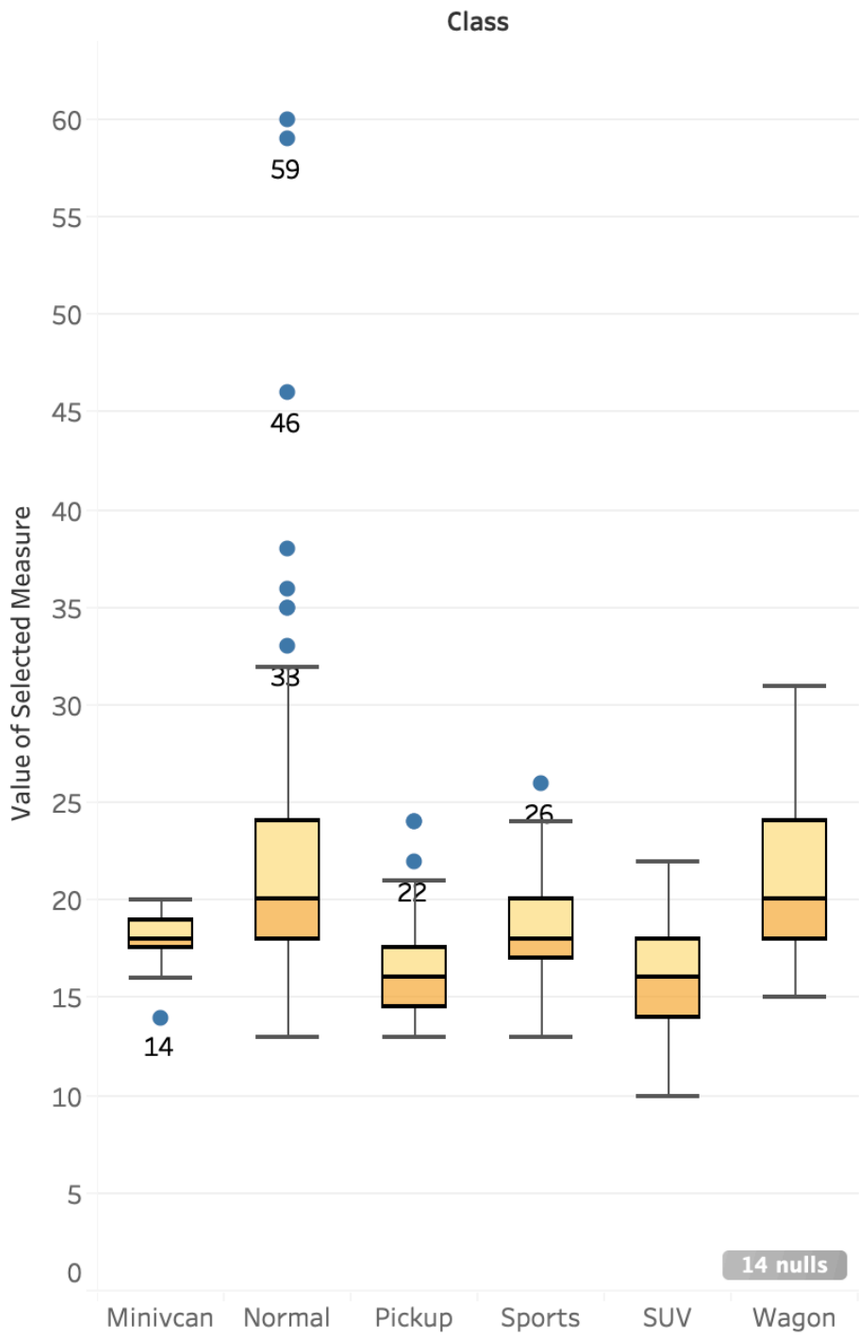
Color | Size | Label | Detail | Tooltip

Vehicle Name

Columns: Class

Rows: Value of Selected Mea..

Box Plot With Parameter . The Selected mesasure: **City MPG**



Selected Measure

- (•) City MPG
- ( ) Dealer Cost
- ( ) Retail Price

Pages

Columns: Class

Rows: Value of Selected Mea..

Filters

Marks

Circle

Color | Size | Label | Detail | Tooltip

Vehicle Name

Box Plot With Parameter

Cla

Value of Selected Measure

60 55 50 45 40 35 30 25 20 15 10 5 0

59 46 38 33 22 14

Minivcan Normal Pickup

Edit Reference Line, Band, or Box

Line | Band | Distribution | Box Plot

Plot Options

Whiskers extend to: Data within 1.5 times the IQR

Hide underlying marks (except outliers)

Formatting

Style: Classic with Dual Fill

Fill: Yellow

Border:

Whiskers:

OK

Selected Measure

- City MPG
- Dealer Cost
- Retail Price

Pages

Filters

Marks

Circle

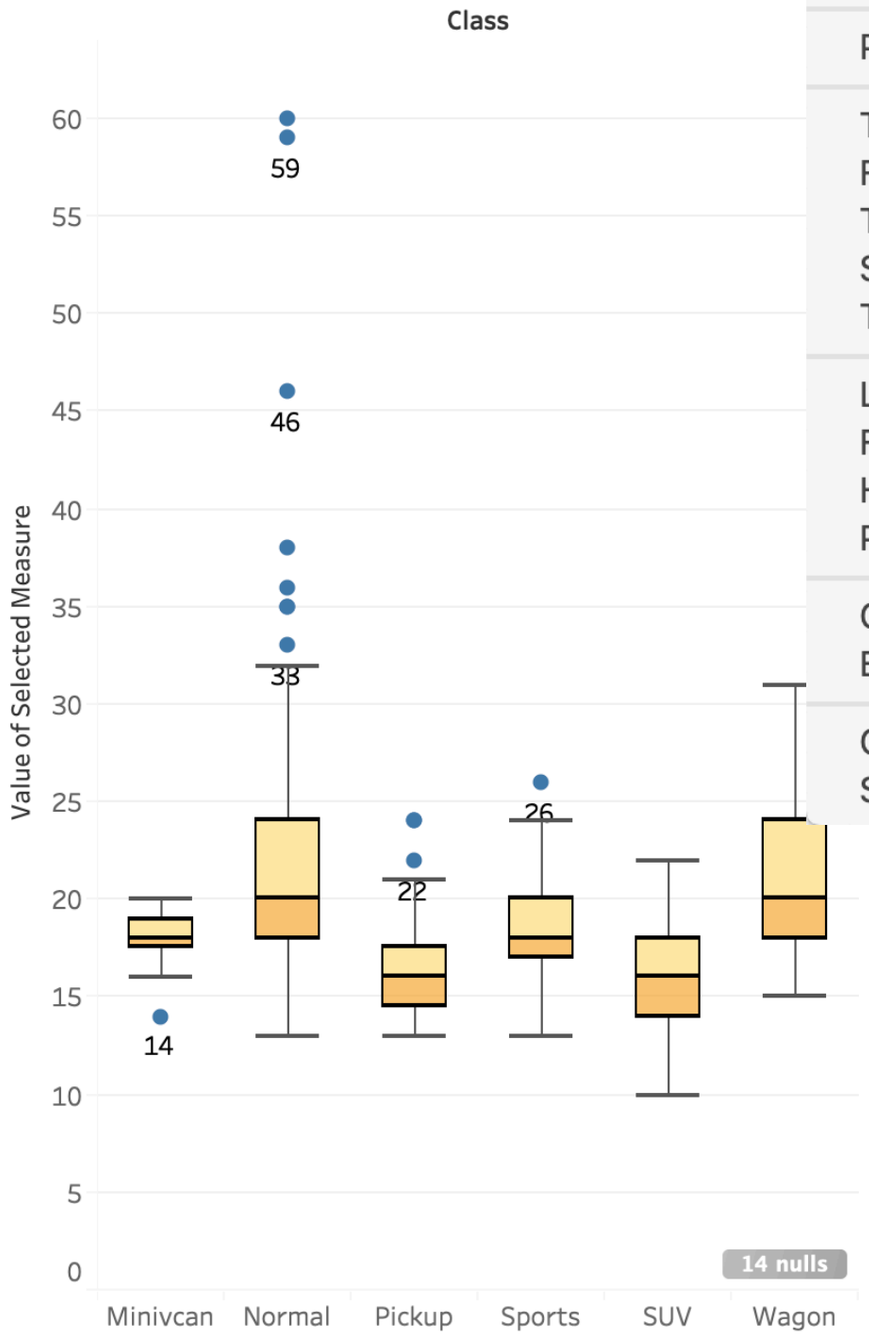
Color | Size | Label | Detail | Tooltip

Vehicle Name

Columns: Class

Rows: Value of Selected Mea..

Box Plot With Parameter . The Selected m



Selected Measure

- (•) City MPG
- ( ) Dealer Cost
- ( ) Retail Price

Analysis | Map | Format | Server

- ✓ Show Mark Labels
- Aggregate Measures
- Stack Marks ▶
- View Data...
- Reveal Hidden Data
- Percentage Of ▶
- Totals ▶
- Forecast ▶
- Trend Lines ▶
- Special Values ▶
- Table Layout ▶
- Legends ▶
- Filters ▶
- Highlighters ▶
- Parameters ▶
- Create Calculated Field...
- Edit Calculated Field ▶
- Cycle Fields
- Swap Rows and Columns ^⌘W

Pages

Filters

Marks

Circle

Color | Size | Label | Detail | Tooltip

Vehicle Name

Columns: Class

Rows: Value of Selected Mea..

Box Plot Wit

Edit Parameter [Selected Measure]

Name: Selected Measure

Comment >>

Properties

Data type: String

Current value: City MPG

Display format:

Allowable values: All | List | Range

List of values

Value	Display As
City MPG	City MPG
Dealer Cost	Dealer Cost
Retail Price	Retail Price
Add	

Add from Parameter ▶

Add from Field ▶

Paste from Clipboard

Clear All

Cancel | OK

Selected Measure

- (•) City MPG
- ( ) Dealer Cost
- ( ) Retail Price

Pages

Filters

Marks

Circle

Color | Size | Label | Detail | Tooltip

Vehicle Name

Columns: Class

Rows: Value of Selected Mea..

Box Plot With Parameter . The Selected mesasure: **City MPG**

Class

60

Selected Measure

- City MPG
- Dealer Cost
- Retail Price

Value of Selected Measure

```
IF [Selected Measure] = "City MPG" THEN [City MPG]
ELSEIF [Selected Measure] = "Dealer Cost" THEN [Dealer Cost]
ELSE [Retail Price]
END
```

The calculation is valid.

1 Dependency | Apply | OK

Minivan Normal Pickup Sports SUV Wagon

Hands On: Cars Dataset

- **Steps:**
  - **One can ask many questions based on this data set. For example:**
    - what are the relationships between the weight and the length? or the price?
    - Is there a correlation between the vehicle model and the MPG (or Lp100K), i.e. specific models have specific range of MPG (or Lp100K)?
    - Does the dealer price mean that we have a better performing vehicle (higher price implies better MPG)?
  - **Try to answer these questions and propose new ones.**

Table Calculations

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

SUM(Numb.. △)

Columns: Drive Train

Rows: Class

Class by Drive Train - % Table

	Drive Train			
Class	AWD	FWD	RWD	Grand To..
Minivcan	0,70% [1]	3,74% [2]	0,23% [3]	**4,67% [1]**
Normal	5,84% [4]	38,79% [5]	12,62% [6]	**57,24% [2]**
Pickup	2,80% [7]		2,80% [8]	**5,61% [3]**
Sports	1,17% [9]	1,87% [10]	8,41% [11]	**11,45% [4]**
SUV	8,88% [12]	5,14% [13]		**14,02% [5]**
Wagon	2,10% [14]	3,27% [15]	1,64% [16]	**7,01% [6]**
Grand Total	**21,50% [1]**	**52,80% [2]**	**25,70% [3]**	**100,00% [1]**

Table Calculation
% of Total Number of Records

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

Table (across)
Table (down)
Table
Cell
Specific Dimensions

Class
Drive Train

At the level

Show calculation assistance

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

SUM(Numb.. △)

Columns: Drive Train

Rows: Class

Class by Drive Train - % Table

	Drive Train			
Class	AWD	FWD	RWD	Grand To..
Minivcan	15,00% [1]	80,00% [2]	5,00% [3]	**100,00% [1]**
Normal	10,20% [1]	67,76% [2]	22,04% [3]	100,00% [1]
Pickup	50,00% [1]	[2]	50,00% [3]	100,00% [1]
Sports	10,20% [1]	16,33% [2]	73,47% [3]	100,00% [1]
SUV	63,33% [1]	36,67% [2]	[3]	100,00% [1]
Wagon	30,00% [1]	46,67% [2]	23,33% [3]	100,00% [1]
Grand Total	21,50% [1]	52,80% [2]	25,70% [3]	100,00% [1]

Table Calculation
% of Total Number of Records

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Cell
- **Specific Dimensions**

- [x] Drive Train
- [ ] Class

At the level

[x] Show calculation assistance

Pages

Filters

Marks

Automatic

Color | Size | Label | Detail | Tooltip

SUM(Numb.. △)

Columns: Drive Train

Rows: Class

Class by Drive Train - % Table

	Drive Train			
Class	AWD	FWD	RWD	Grand To..
Minivcan	3,26% [1]	7,08% [1]	0,91% [1]	4,67% [1]
Normal	27,17% [2]	73,45% [2]	49,09% [2]	57,24% [2]
Pickup	13,04% [3]	[3]	10,91% [3]	5,61% [3]
Sports	5,43% [4]	3,54% [4]	32,73% [4]	11,45% [4]
SUV	41,30% [5]	9,73% [5]	[5]	14,02% [5]
Wagon	9,78% [6]	6,19% [6]	6,36% [6]	7,01% [6]
Grand Total	100,00% [1]	100,00% [1]	100,00% [1]	100,00% [1]

Table Calculation
% of Total Number of Records

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

Table (across)
Table (down)
Table
Cell
Specific Dimensions

Class
Drive Train

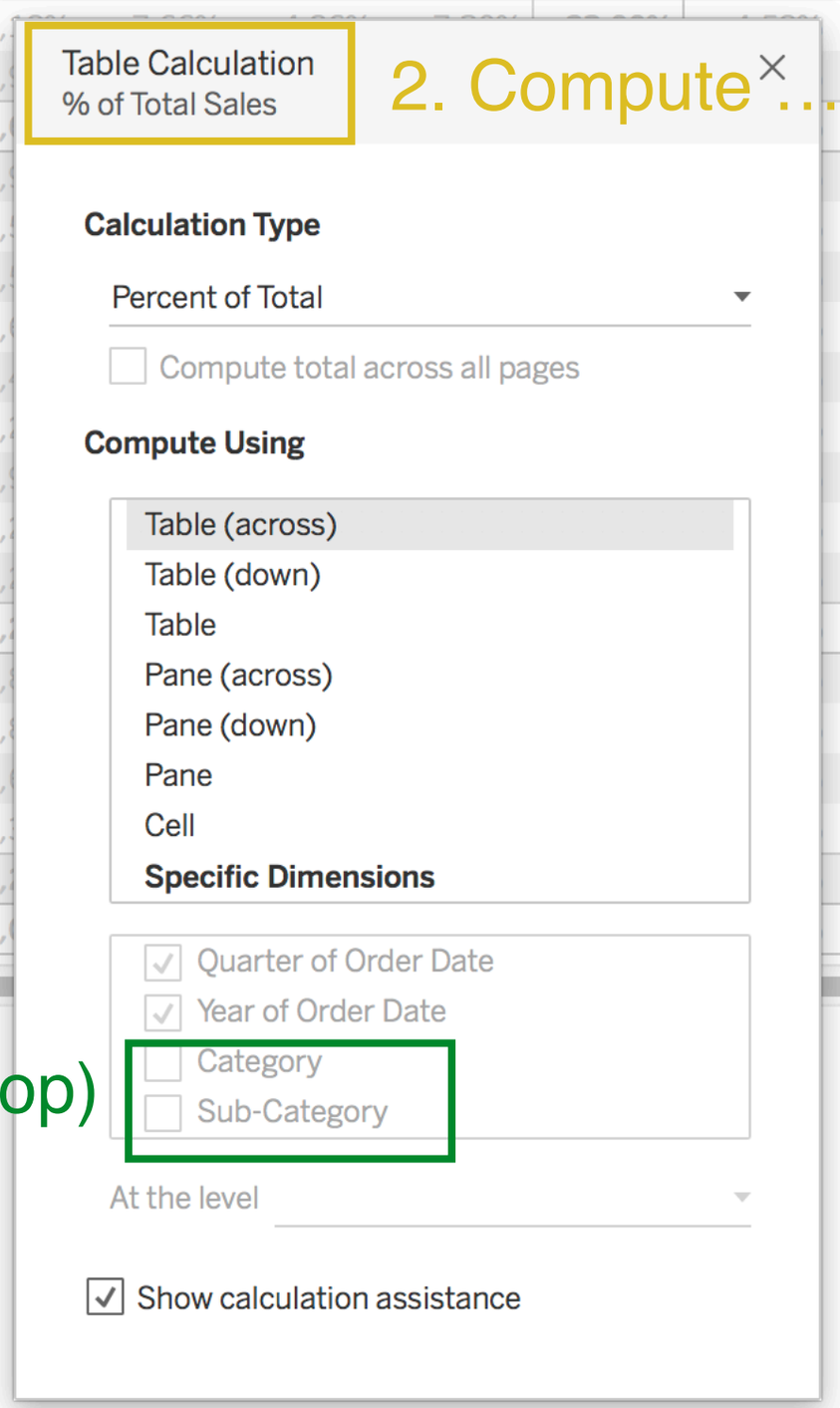
At the level

Show calculation assistance

Sheet 1

		Order Date										
		2014					2015					
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q
Furniture	Bookcases	2,36%	1,49%	4,07%	9,52%	**17,44%**	4,10%	2,85%	15,10%	11,49%	**33,55%**	3,11%	3,62
	Chairs	3,07%	4,06%	6,94%	9,44%	**23,52%**	2,88%	4,23%	5,56%	9,17%	**21,84%**	2,85%	5,78
	Furnishings	1,73%	2,40%	5,30%	5,65%	**15,08%**	3,						5,75
	Tables	3,99%	5,23%	4,65%	8,40%	**22,27%**	4,						6,28
	Total	**3,05%**	**3,78%**	**5,65%**	**8,69%**	21,19%	3,						**5,58**
Office Supplies	Appliances	0,84%	3,47%	4,27%	5,66%	**14,24%**	1,						4,02
	Art	2,45%	5,69%	4,50%	9,71%	**22,34%**	2,						5,41
	Binders	1,32%	4,74%	9,80%	5,53%	**21,38%**	5,						2,30
	Envelopes	3,32%	4,36%	4,38%	11,34%	**23,40%**	3,						2,53
	Fasteners	3,23%	5,28%	6,80%	6,56%	**21,87%**	2,						5,00
	Labels	2,04%	5,96%	4,95%	9,81%	**22,76%**	6,						4,58
	Paper	2,01%	2,80%	5,43%	8,67%	**18,90%**	2,						8,55
	Storage	3,37%	3,68%	6,07%	9,36%	**22,48%**	2,						6,17
	Supplies	0,55%	9,17%	18,84%	2,27%	**30,84%**	0,						5,24
	Total	**2,02%**	**4,35%**	**7,50%**	**7,24%**	21,11%	3,						**4,81**
Technology	Accessories	1,83%	2,34%	4,60%	6,18%	**14,94%**	1,						4,32
	Copiers		2,01%	2,75%	2,50%	**7,26%**	2,						6,98
	Machines	13,38%	0,52%	11,85%	7,03%	**32,78%**	1,						12,83
	Phones	2,69%	5,86%	4,10%	10,80%	**23,45%**	2,						5,50
	Total	**4,46%**	**3,26%**	**5,71%**	**7,54%**	20,96%	2,						**7,19**
Grand Total		**3,24%**	**3,77%**	**6,25%**	**7,82%**	21,08%	3,						**5,92**

1. For each (bottom to top)



Sheet 1

		Order Date										
		2014					2015					
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1
Furniture	Bookcases	3,65%	1,97%	3,26%	6,09%	4,14%	6,85%	3,68%	13,32%	7,24%	8,19%	3,83%
	Chairs	13,55%	15,42%	15,88%	17,26%	15,95%	13,72%	15,59%	14,03%	16,52%	15,25%	10,03%
	Furnishings	2,13%	2,54%	3,38%	2,89%	2,86%						
	Tables	11,10%	12,50%	6,69%	9,68%	9,52%						
	Total	30,43%	32,43%	29,21%	35,92%	32,46%						
Office Supplies	Appliances	1,22%	4,31%	3,20%	3,39%	3,16%						
	Art	0,89%	1,78%	0,85%	1,47%	1,25%						
	Binders	3,60%	11,13%	13,88%	6,26%	8,98%						
	Envelopes	0,74%	0,83%	0,50%	1,04%	0,80%						
	Fasteners	0,13%	0,18%	0,14%	0,11%	0,14%						
	Labels	0,34%	0,86%	0,43%	0,68%	0,59%						
	Paper	2,12%	2,54%	2,97%	3,79%	3,06%						
	Storage	10,13%	9,52%	9,46%	11,67%	10,39%						
	Supplies	0,35%	4,95%	6,12%	0,59%	2,97%						
	Total	19,52%	36,10%	37,54%	28,99%	31,34%						
Technology	Accessories	4,11%	4,52%	5,37%	5,76%	5,17%						
	Copiers		3,47%	2,86%	2,08%	2,24%						
	Machines	34,00%	1,13%	15,61%	7,41%	12,81%						
	Phones	11,94%	22,35%	9,41%	19,84%	15,98%						
	Total	50,05%	31,47%	33,25%	35,09%	36,20%						
Grand Total		100,00%	100,00%	100,00%	100,00%	100,00%						

Table Calculation
% of Total Sales

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Pane (across)
- Pane (down)
- Pane
- Cell
- **Specific Dimensions**

- [x] Category
- [x] Sub-Category
- [ ] Quarter of Order Date
- [ ] Year of Order Date

At the level

[x] Show calculation assistance

Sheet 1

		Order Date											
		2014					2015						
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q2
Furniture	Bookcases	0,12%	0,07%	0,20%	0,48%	**0,87%**	0,21%	0,14%	0,76%	0,57%	**1,68%**	0,16%	0,18%
	Chairs	0,44%	0,58%	0,99%	1,35%	**3,36%**	0,41%	0,60%	0,80%	1,31%	**3,12%**	0,41%	0,83%
	Furnishings	0,07%	0,10%	0,21%	0,23%	**0,60%**	0,[illegible]						0,23%
	Tables	0,36%	0,47%	0,42%	0,76%	**2,01%**	0,[illegible]						0,57%
	Total	**0,99%**	**1,22%**	**1,83%**	**2,81%**	**6,84%**	**1,[illegible]**						**1,80%**
Office Supplies	Appliances	0,04%	0,16%	0,20%	0,27%	**0,67%**	0,[illegible]						0,19%
	Art	0,03%	0,07%	0,05%	0,11%	**0,26%**	0,[illegible]						0,06%
	Binders	0,12%	0,42%	0,87%	0,49%	**1,89%**	0,[illegible]						0,20%
	Envelopes	0,02%	0,03%	0,03%	0,08%	**0,17%**	0,[illegible]						0,02%
	Fasteners	0,00%	0,01%	0,01%	0,01%	**0,03%**	0,[illegible]						0,01%
	Labels	0,01%	0,03%	0,03%	0,05%	**0,12%**	0,[illegible]						0,02%
	Paper	0,07%	0,10%	0,19%	0,30%	**0,65%**	0,[illegible]						0,29%
	Storage	0,33%	0,36%	0,59%	0,91%	**2,19%**	0,[illegible]						0,60%
	Supplies	0,01%	0,19%	0,38%	0,05%	**0,63%**	0,[illegible]						0,11%
	Total	**0,63%**	**1,36%**	**2,35%**	**2,27%**	**6,61%**	**1,[illegible]**						**1,51%**
Technology	Accessories	0,13%	0,17%	0,34%	0,45%	**1,09%**	0,[illegible]						0,31%
	Copiers		0,13%	0,18%	0,16%	**0,47%**	0,[illegible]						0,45%
	Machines	1,10%	0,04%	0,98%	0,58%	**2,70%**	0,[illegible]						1,06%
	Phones	0,39%	0,84%	0,59%	1,55%	**3,37%**	0,[illegible]						0,79%
	Total	**1,62%**	**1,19%**	**2,08%**	**2,74%**	**7,63%**	**0,[illegible]**						**2,62%**
Grand Total		**3,24%**	**3,77%**	**6,25%**	**7,82%**	**21,08%**	**3,[illegible]**						**5,92%**

Table Calculation
% of Total Sales

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Pane (across)
- Pane (down)
- Pane
- Cell
- **Specific Dimensions**

- Category
- Sub-Category
- Quarter of Order Date
- Year of Order Date

At the level

Show calculation assistance

Sheet 1

		Order Date											
		2014					2015						
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q2
Furniture	Bookcases	13,56%	8,53%	23,34%	54,58%	100,00%	12,23%	8,50%	45,02%	34,26%	100,00%	13,59%	15,81%
	Chairs	13,06%	17,27%	29,52%	40,14%	100,00%	13,17%	19,37%	25,47%	41,99%	100,00%	11,14%	22,62%
	Furnishings	11,46%	15,89%	35,16%	37,49%	100,00%	13,						18,90%
	Tables	17,93%	23,47%	20,86%	37,74%	100,00%	26,						21,37%
	Total	14,41%	17,85%	26,69%	41,04%	100,00%	16,						20,82%
Office Supplies	Appliances	5,91%	24,33%	29,98%	39,77%	100,00%	9,						16,61%
	Art	10,96%	25,46%	20,13%	43,45%	100,00%	11,						24,63%
	Binders	6,16%	22,15%	45,83%	25,86%	100,00%	30,						9,44%
	Envelopes	14,20%	18,62%	18,71%	48,46%	100,00%	13,						8,81%
	Fasteners	14,77%	24,13%	31,10%	30,00%	100,00%	13,						15,75%
	Labels	8,98%	26,19%	21,74%	43,09%	100,00%	26,						20,24%
	Paper	10,66%	14,79%	28,71%	45,84%	100,00%	15,						32,48%
	Storage	14,98%	16,38%	26,99%	41,65%	100,00%	11,						23,48%
	Supplies	1,79%	29,75%	61,09%	7,37%	100,00%	6,						17,13%
	Total	9,57%	20,59%	35,53%	34,31%	100,00%	16,						18,80%
Technology	Accessories	12,22%	15,63%	30,81%	41,34%	100,00%	7,						17,24%
	Copiers		27,65%	37,88%	34,47%	100,00%	16,						21,05%
	Machines	40,81%	1,58%	36,15%	21,46%	100,00%	11,						43,42%
	Phones	11,49%	24,99%	17,46%	46,05%	100,00%	11,						23,00%
	Total	21,26%	15,54%	27,24%	35,96%	100,00%	11,						26,55%
Grand Total		15,37%	17,87%	29,66%	37,09%	100,00%	14,						22,34%

Table Calculation
% of Total Sales

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Pane (across)
- Pane (down)
- Pane
- Cell
- **Specific Dimensions**

- [x] Quarter of Order Date
- [ ] Category
- [ ] Sub-Category
- [ ] Year of Order Date

At the level

[x] Show calculation assistance

Sheet 1

		Order Date											
		2014					2015					201	
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q2
Furniture	Bookcases	11,99%	6,09%	11,15%	16,95%	12,75%	17,22%	11,88%	34,99%	20,01%	22,60%	14,67%	10,03%
	Chairs	44,54%	47,54%	54,35%	48,06%	49,14%	34,51%	50,40%	36,85%	45,65%	42,07%	38,39%	45,84%
	Furnishings	7,00%	7,83%	11,59%	8,03%	8,80%	10,						12,73%
	Tables	36,47%	38,54%	22,92%	26,96%	29,32%	37,						31,40%
	Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,					100,00%	
Office Supplies	Appliances	6,23%	11,93%	8,52%	11,69%	10,09%	9,						12,51%
	Art	4,57%	4,94%	2,26%	5,05%	3,99%	3,						4,25%
	Binders	18,44%	30,83%	36,96%	21,59%	28,65%	48,						13,55%
	Envelopes	3,77%	2,30%	1,34%	3,59%	2,54%	2,						1,21%
	Fasteners	0,67%	0,51%	0,38%	0,38%	0,44%	0,						0,44%
	Labels	1,76%	2,38%	1,15%	2,35%	1,87%	3,						1,65%
	Paper	10,88%	7,02%	7,90%	13,06%	9,77%	10,						19,40%
	Storage	51,91%	26,38%	25,19%	40,24%	33,16%	22,						39,92%
	Supplies	1,77%	13,71%	16,31%	2,04%	9,48%	0,						7,07%
	Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,					100,00%	
Technology	Accessories	8,21%	14,35%	16,14%	16,41%	14,27%	17,						12,02%
	Copiers		11,02%	8,61%	5,93%	6,19%	23,						17,37%
	Machines	67,94%	3,60%	46,95%	21,11%	35,39%	16,						40,39%
	Phones	23,86%	71,03%	28,30%	56,55%	44,15%	42,						30,22%
	Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,					100,00%	
Grand Total		100,00%	100,00%	100,00%	100,00%	100,00%	100,					100,00%	

Table Calculation
% of Total Sales

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Pane (across)
- Pane (down)
- Pane
- Cell
- **Specific Dimensions**

- [x] Sub-Category
- [ ] Category
- [ ] Quarter of Order Date
- [ ] Year of Order Date

At the level

[x] Show calculation assistance

Sheet 1

		Order Date												
		2014					2015							201
Category	Sub-Catego..	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q2	
Furniture	Accessories													
	Appliances													
	Art													
	Binders													
	Bookcases	1,73%	1,09%	2,97%	6,96%	**12,75%**	2,7						2,09%	4,9
	Chairs	6,42%	8,49%	14,51%	19,72%	**49,14%**	5,5						9,54%	10,8
	Copiers													
	Envelopes													
	Fasteners													
	Furnishings	1,01%	1,40%	3,09%	3,30%	**8,80%**	1,7						2,65%	3,1
	Labels													
	Machines													
	Paper													
	Phones													
	Storage													
	Supplies													
	Tables	5,26%	6,88%	6,12%	11,06%	**29,32%**	6,(						6,54%	7,6
	Total	**14,41%**	**17,85%**	**26,69%**	**41,04%**	**100,00%**	16,(						20,82%	26,5
Office Supplies	Accessories													
	Appliances	0,60%	2,46%	3,03%	4,01%	**10,09%**	1,5						2,35%	4,2
	Art	0,44%	1,02%	0,80%	1,73%	**3,99%**	0,5						0,80%	1,(
	Binders	1,77%	6,35%	13,13%	7,41%	**28,65%**	8,2						2,55%	6,8
	Bookcases													
	Chairs													
	Copiers													
	Envelopes	0,36%	0,47%	0,48%	1,23%	**2,54%**	0,4						0,23%	0,8
	Fasteners	0,06%	0,11%	0,14%	0,13%	**0,44%**	0,(						0,08%	0,1
	Furnishings													
	Labels	0,17%	0,49%	0,41%	0,81%	**1,87%**	0,5						0,31%	0,4
	Machines													
	Paper	1,04%	1,45%	2,81%	4,48%	**9,77%**	1,69%	1,77%	2,94%	4,74%	**11,14%**	1,49%	3,65%	2,7
	Phones													
	Storage	4,97%	5,43%	8,95%	13,81%	**33,16%**	3,70%	6,28%	12,61%	10,23%	**32,83%**	4,81%	7,51%	7,8
	Supplies	0,17%	2,82%	5,79%	0,70%	**9,48%**	0,09%	0,39%	0,77%	0,18%	**1,42%**	4,84%	1,33%	0,4

Table Calculation
% of Total Sales

Calculation Type

Percent of Total

Compute total across all pages

Compute Using

- Table (across)
- Table (down)
- Table
- Pane (across)
- Pane (down)
- Pane
- Cell
- **Specific Dimensions**

- [x] Sub-Category
- [x] Quarter of Order Date
- [ ] Category
- [ ] Year of Order Date

At the level

[x] Show calculation assistance

Table Calculations

- **See more here**

The following quick table calculations are available in Tableau for you to use:

- Running total
- Difference
- Percent difference
- Percent of total
- Rank
- Percentile
- Moving average
- YTD total
- Compound growth rate
- Year of year growth
- YTD growth

https://onlinehelp.tableau.com/current/pro/desktop/en-us/calculations_tablecalculations_definebasic_runningtotal.htm

Hands On: Cars Dataset

- **Steps:**
  - **Free browsing:**
    - Select a subset of data (for instance grouping some brands by nationality or geography or by class) that contains an obvious correlation (exploratory visualization).
    - State a hypotheses and confirm it using the full data set (confirmatory visualization)
    - Present your result using some slides (presentation visualization)

THANK YOU!

jmp@fct.unl.pt

fpb@fct.unl.pt

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

NOVALINCS
LABORATORY FOR COMPUTER SCIENCE AND INFORMATICS

Departamento de Informática, FCT/UNL
Quinta da Torre P-2829-516
CAPARICA, Portugal

di.secretariado(AT)fct.unl.pt

(+351) 212948536 (direct)
(+351) 212948300 (central)
(+351) 212948500 (central)