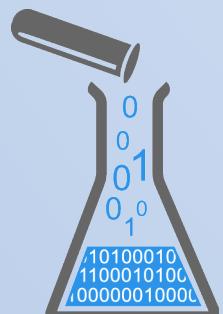
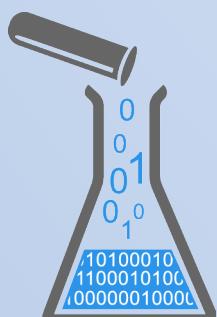


DATA TRANSFORMATION

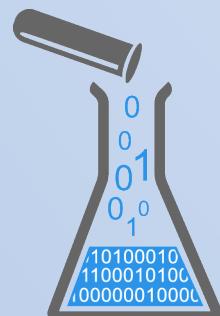
A Framework for Exploratory Data Analysis



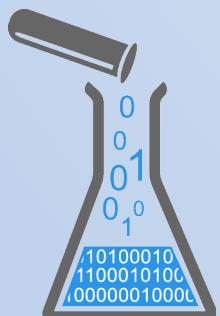
How many of you consider
yourselves data scientists?



How many of you spend a
lot of time **exploring data?**



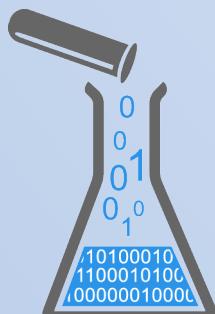
How many of you have a
formal process for
data exploration?



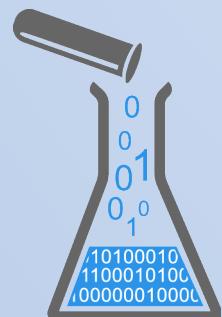
ABOUT ME – TONY OJEDA



Founder of District Data Labs
Research lab and open source collaborative
Business & finance background
Self-taught programmer (R & Python)

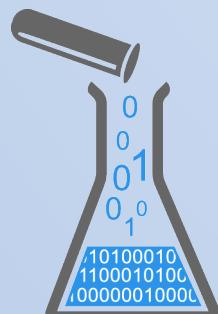


HOW I THINK ABOUT DATA



ORGANIZATION

Categories, Classifications, Taxonomies, Tags, etc.



EXPLORATION FRAMEWORK

Prep Phase

Identify

Types of Information

Entities in Data Set

Review

Transformation Methods

Visualization Methods

Create

Category Aggregations

Continuous Bins

Cluster Categories

Explore Phase

Filter + Aggregate

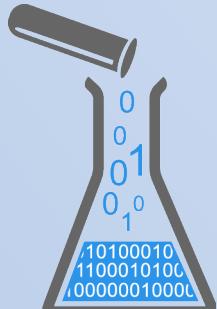
Field Relationships

Entity Relationships

Visualization

Over Time

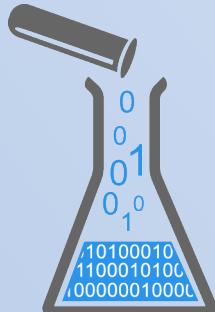
Insights



THE DATA: EPA VEHICLE FUEL ECONOMY

make	model	year	displ	cylinders	trany	drive	VClass	fuelType	barrels08	city08	highway08	comb08	co2TailpipeGpm	fuelCost08
AM General	DJ Po Vehicle 2WD	1984	2.5	4	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	19.388824	18	17	17	522.764706	2000
AM General	FJ8c Post Office	1984	4.2	6	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
AM General	Post Office DJ5 2WD	1985	2.5	4	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	20.600625	16	17	16	555.4375	2100
AM General	Post Office DJ8 2WD	1985	4.2	6	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
ASC Incorporated	GNX	1987	3.8	6	Automatic 4-spd	Rear-Wheel Drive	Midsize Cars	Premium	20.600625	14	21	16	555.4375	2550
Acura	2.2CL/3.0CL	1997	2.2	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	26	22	403.954545	1550
Acura	2.2CL/3.0CL	1997	2.2	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	22	28	24	370.291667	1400
Acura	2.2CL/3.0CL	1997	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	18	26	20	444.35	1700
Acura	2.3CL/3.0CL	1998	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	19	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1998	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1998	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.3CL/3.0CL	1999	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1999	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1999	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.5TL	1995	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1997	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1997	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1998	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	17	23	20	444.35	2050

<http://www.fueleconomy.gov/feg/epadata/vehicles.csv.zip>



IDENTIFY

Identify

Types of Information

Entities in Data Set

Review

Transformation
Methods

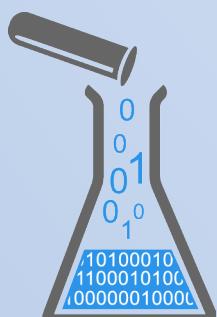
Visualization Methods

Create

Category Aggregations

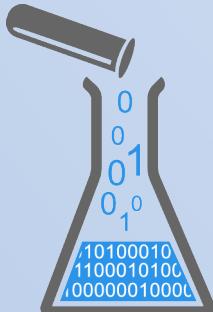
Continuous Bins

Cluster Categories



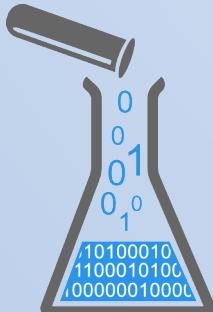
IDENTIFY TYPES OF INFORMATION

make	model	year	displ	cylinders	trany	drive	VClass	fuelType	barrels08	city08	highway08	comb08	co2TailpipeGpm	fuelCost08
AM General	DJ Po Vehicle 2WD	1984	2.5	4	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	19.388824	18	17	17	522.764706	2000
AM General	FJ8c Post Office	1984	4.2	6	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
AM General	Post Office DJ5 2WD	1985	2.5	4	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	20.600625	16	17	16	555.4375	2100
AM General	Post Office DJ8 2WD	1985	4.2	6	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
ASC Incorporated	GNX	1987	3.8	6	Automatic 4-spd	Rear-Wheel Drive	Midsize Cars	Premium	20.600625	14	21	16	555.4375	2550
Acura	2.2CL/3.0CL	1997	2.2	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	26	22	403.954545	1550
Acura	2.2CL/3.0CL	1997	2.2	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	22	28	24	370.291667	1400
Acura	2.2CL/3.0CL	1997	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	18	26	20	444.35	1700
Acura	2.3CL/3.0CL	1998	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	19	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1998	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1998	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.3CL/3.0CL	1999	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1999	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1999	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.5TL	1995	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1997	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1997	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1998	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	17	23	20	444.35	2050

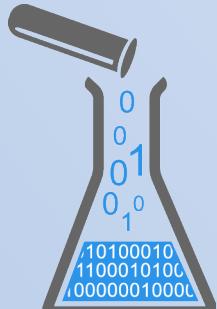
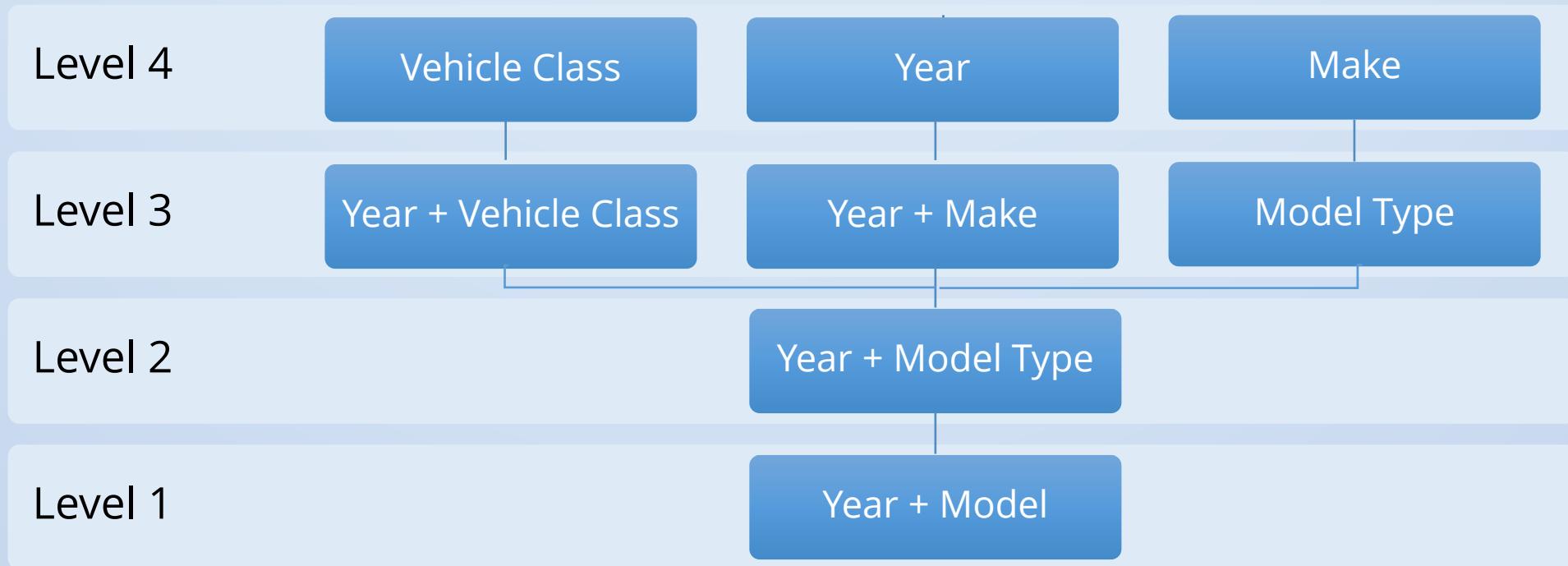


IDENTIFY ENTITIES IN THE DATA

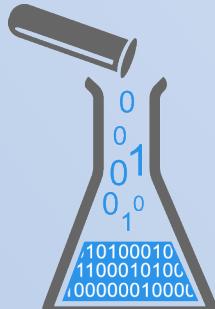
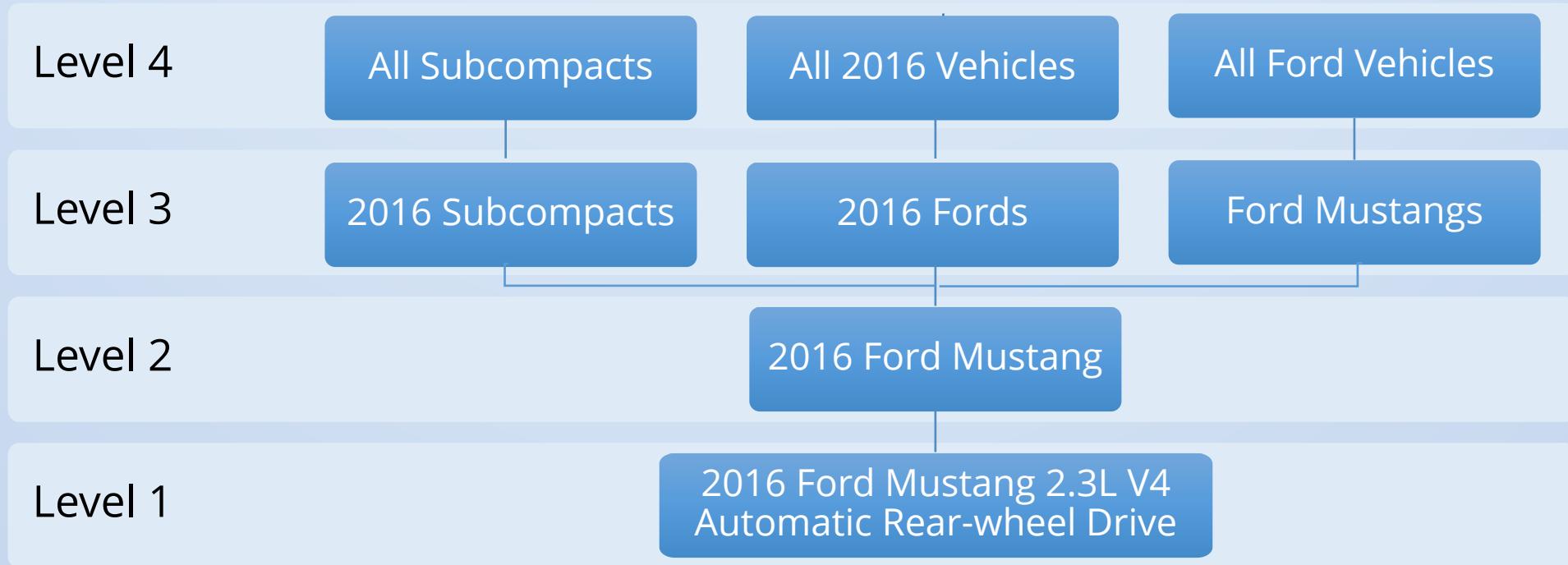
make	model	year	displ	cylinders	trany	drive	VClass	fuelType	barrels08	city08	highway08	comb08	co2TailpipeGpm	fuelCost08
AM General	DJ Po Vehicle 2WD	1984	2.5	4	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	19.388824	18	17	17	522.764706	2000
AM General	FJ8c Post Office	1984	4.2	6	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
AM General	Post Office DJ5 2WD	1985	2.5	4	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	20.600625	16	17	16	555.4375	2100
AM General	Post Office DJ8 2WD	1985	4.2	6	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
ASC Incorporated	GNX	1987	3.8	6	Automatic 4-spd	Rear-Wheel Drive	Midsize Cars	Premium	20.600625	14	21	16	555.4375	2550
Acura	2.2CL/3.0CL	1997	2.2	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	26	22	403.954545	1550
Acura	2.2CL/3.0CL	1997	2.2	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	22	28	24	370.291667	1400
Acura	2.2CL/3.0CL	1997	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	18	26	20	444.35	1700
Acura	2.3CL/3.0CL	1998	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	19	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1998	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1998	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.3CL/3.0CL	1999	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1999	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1999	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.5TL	1995	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1997	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1997	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1998	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	17	23	20	444.35	2050



ENTITIES IN OUR DATA SET



ENTITIES IN OUR DATA SET



REVIEW

Identify

Types of Information

Entities in Data Set

Review

Transformation
Methods

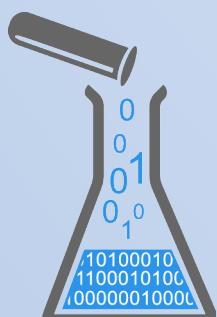
Visualization Methods

Create

Category Aggregations

Continuous Bins

Cluster Categories



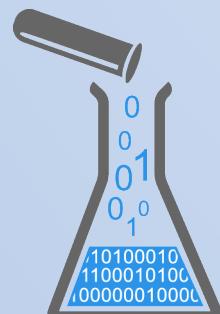
TRANSFORMATION METHODS

Filtering

Aggregation/Disaggregation

Pivoting

Graph Transformation



VISUALIZATION METHODS

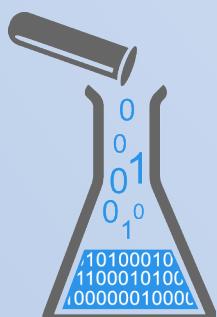
Barcharts 

Multi-line Graphs 

Scatter plots/matrices 

Heatmaps 

Network Visualizations 



CREATE

Identify

Types of Information

Entities in Data Set

Review

Transformation
Methods

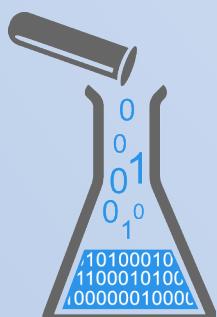
Visualization Methods

Create

Category Aggregations

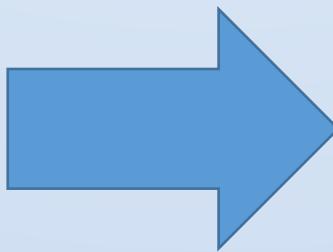
Continuous Bins

Cluster Categories

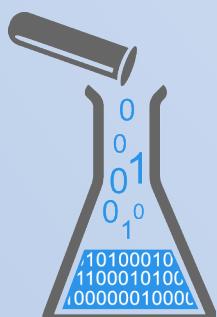


CATEGORY AGGREGATIONS

Transmission
Automatic 3-spd
Automatic 4-spd
Manual 5-spd
Automatic (S5)
Manual 6-spd
Automatic 5-spd
Auto(AM8)
Auto(AV-S7)
Automatic (S6)
Automatic (S9)
Manual 4-spd
+ 33 more

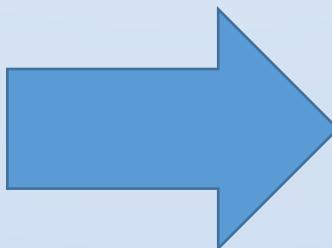


Transmission Type
Automatic
Manual

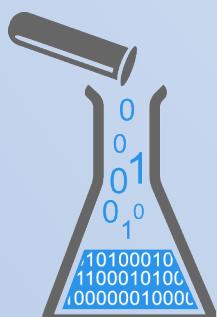


CATEGORY AGGREGATIONS

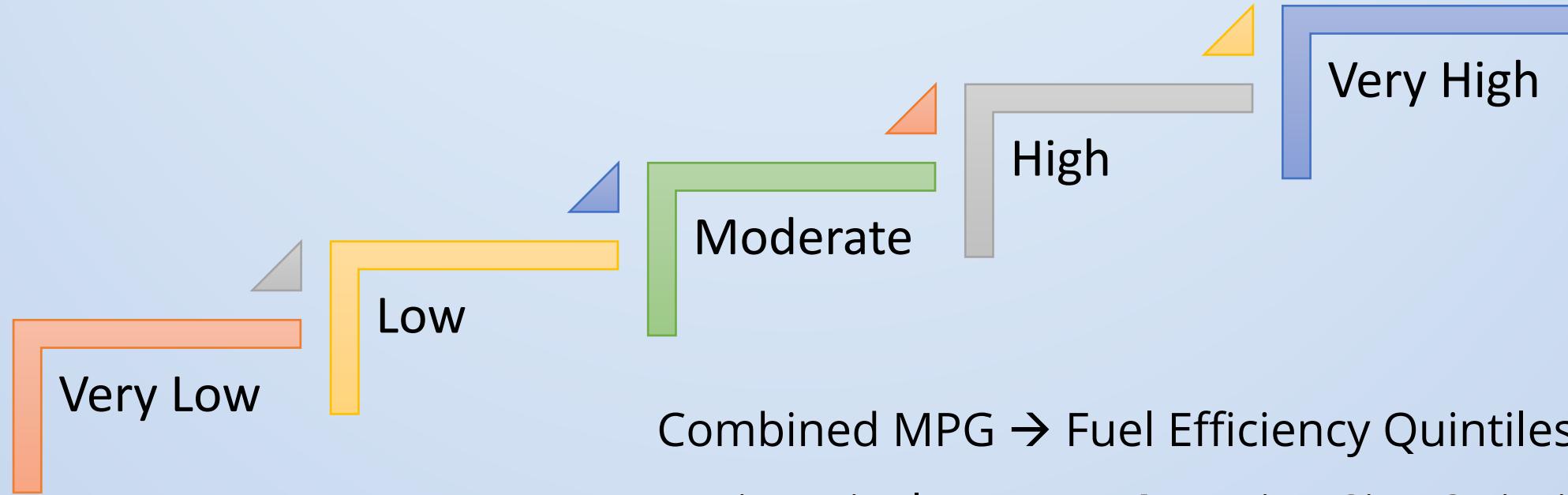
Vehicle Class
Special Purpose Vehicle 2WD
Midsize Cars
Subcompact Cars
Compact Cars
Sport Utility Vehicle - 4WD
Small Sport Utility Vehicle 2WD
Small Sport Utility Vehicle 4WD
Two Seaters
Small Station Wagons
Minicompact Cars
Minivan - 4WD
+ 23 more



Vehicle Category
Small Cars
Midsize Cars
Large Cars
Station Wagons
Pickup Trucks
Special Purpose
Sport Utility
Vans & Minivans



CATEGORIES FROM CONTINUOUS

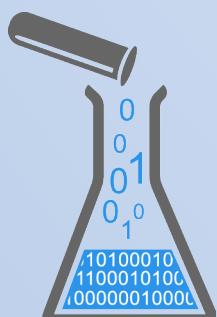


Combined MPG → Fuel Efficiency Quintiles

Engine Displacement → Engine Size Quintiles

CO2 Emission → Emission Quintiles

Fuel Cost → Fuel Cost Quintiles



CLUSTER CATEGORIES

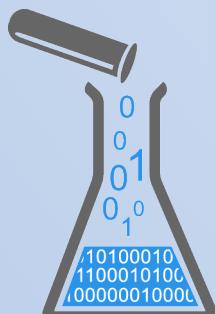
Automatically creates new categories (saves time).

Takes multiple fields into consideration together.

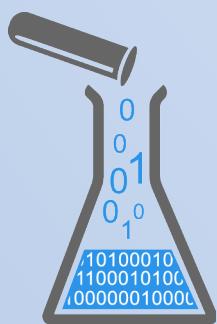
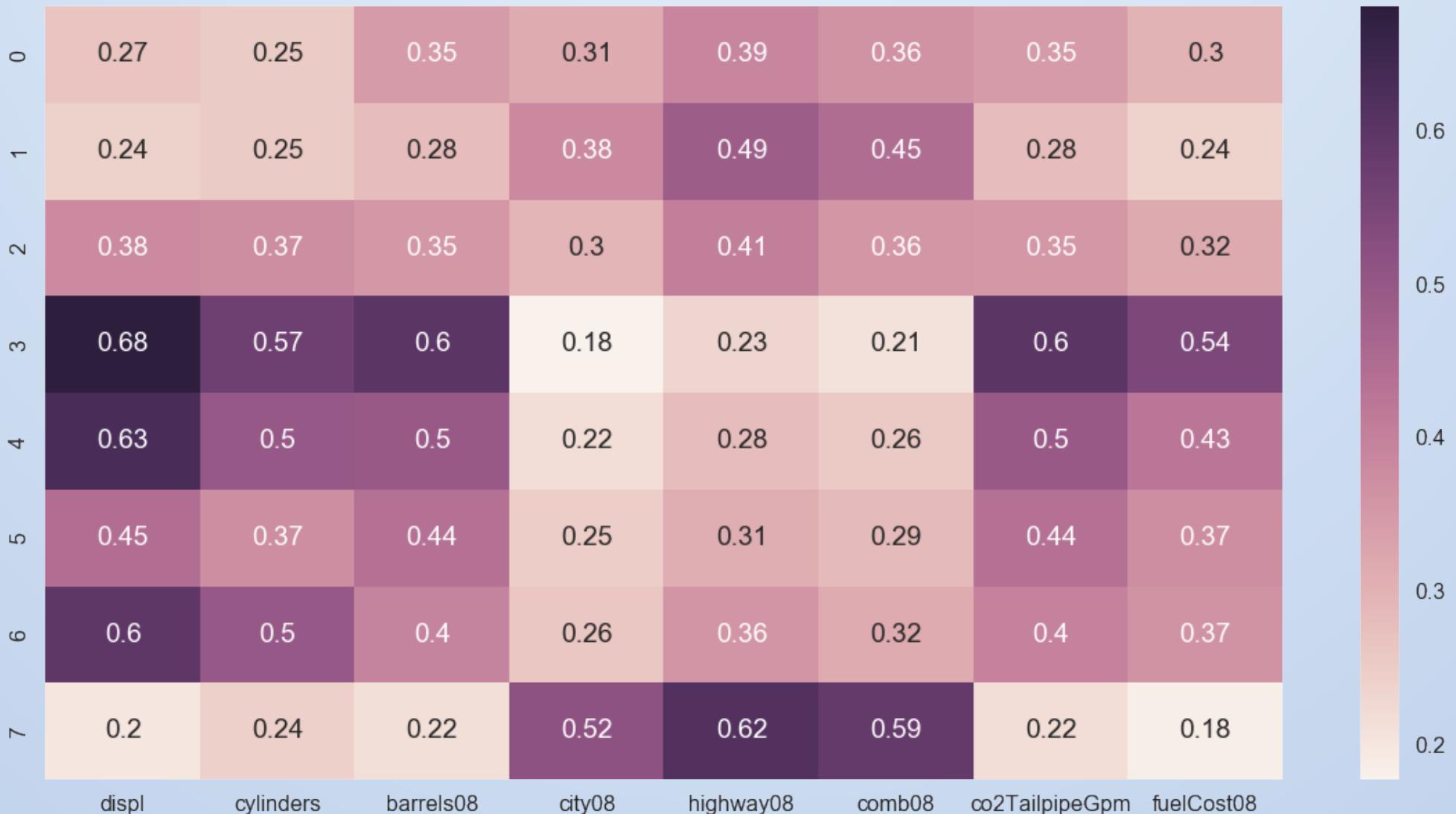
Groups things in ways you may not have thought of.

Number of clusters?
Looking for relatively clear boundaries.

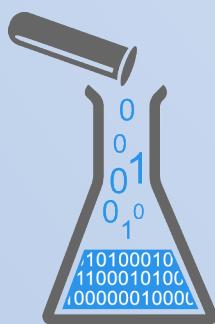
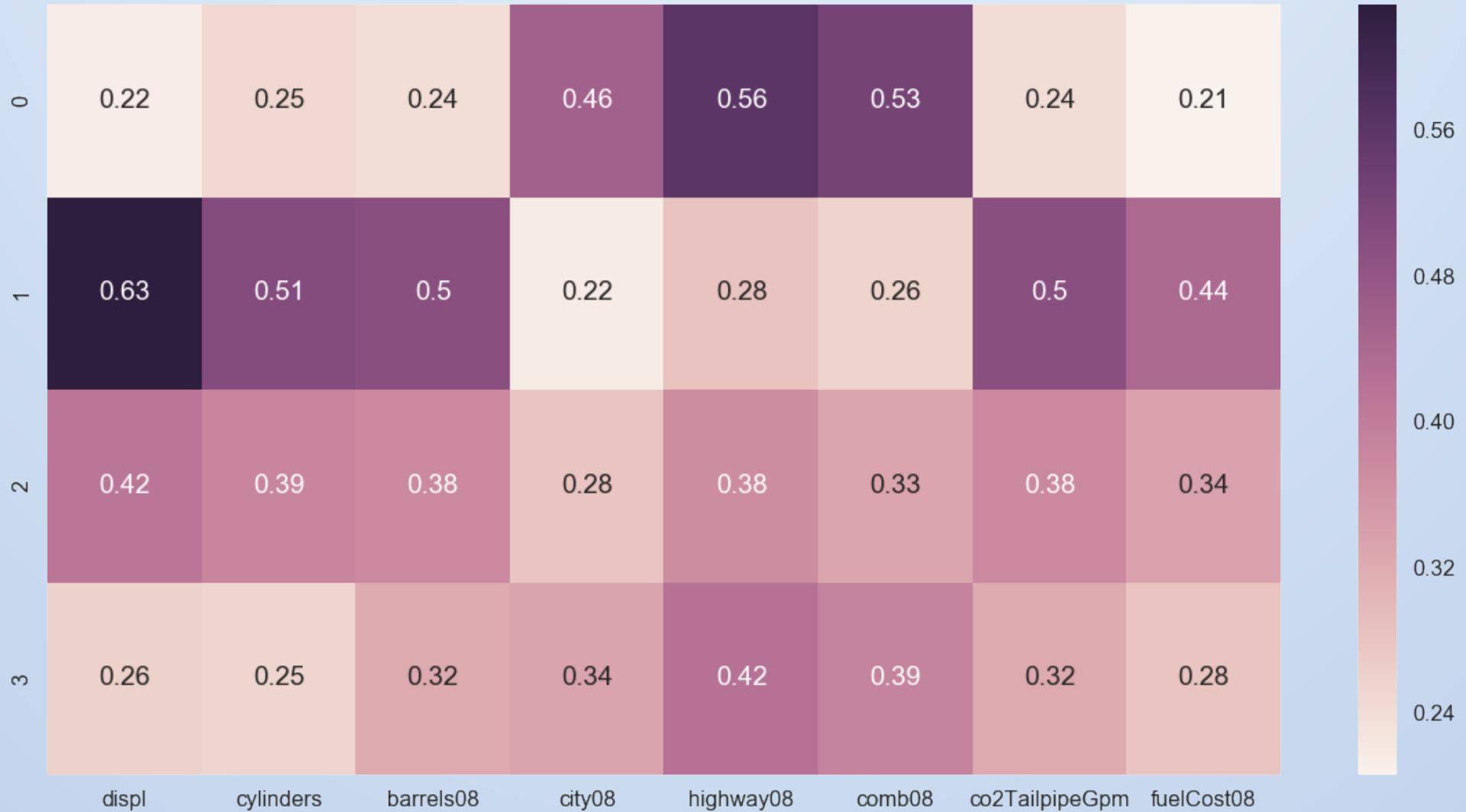
Come up with descriptive names for clusters.



VEHICLE CLUSTERS = 8



VEHICLE CLUSTERS = 4



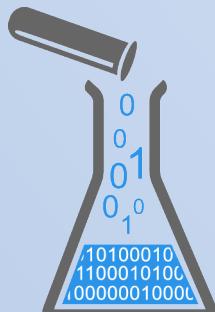
ASSIGN DESCRIPTIVE NAMES

Cluster 0 → Small Very Efficient

Cluster 1 → Large Inefficient

Cluster 2 → Midsized Balanced

Cluster 3 → Small Moderately Efficient



EXPLORE PHASE

Filter + Aggregate

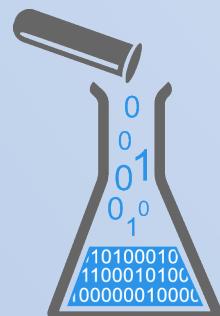
Field Relationships

Entity Relationships

Visualization

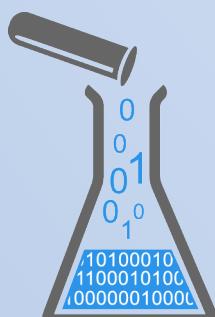
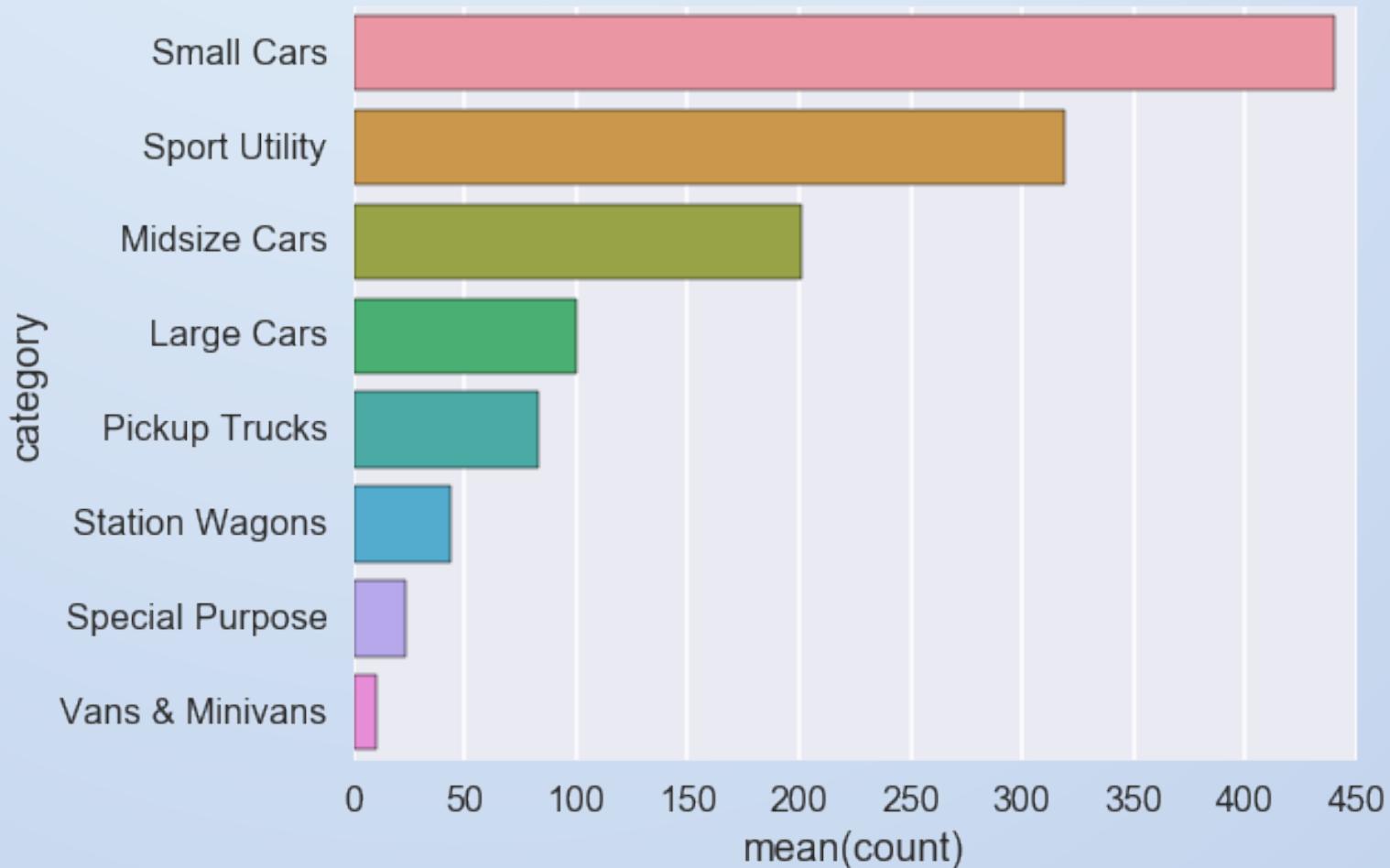
Over Time

Insights

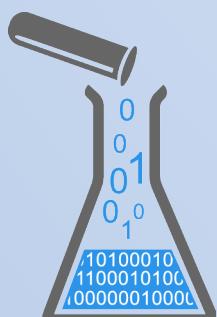
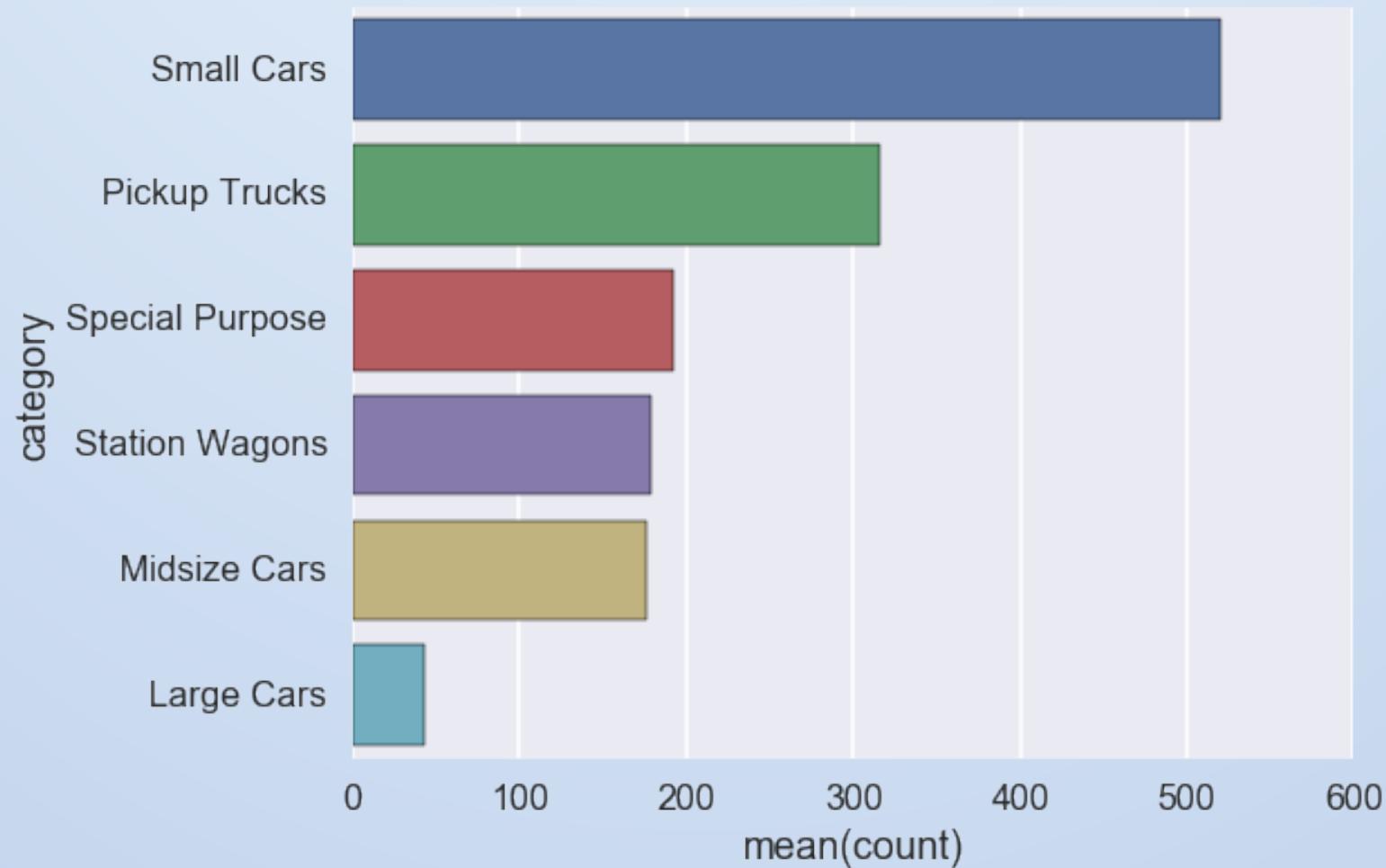


VEHICLE CATEGORY COUNTS (2016)

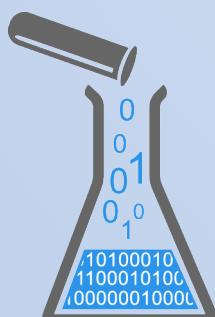
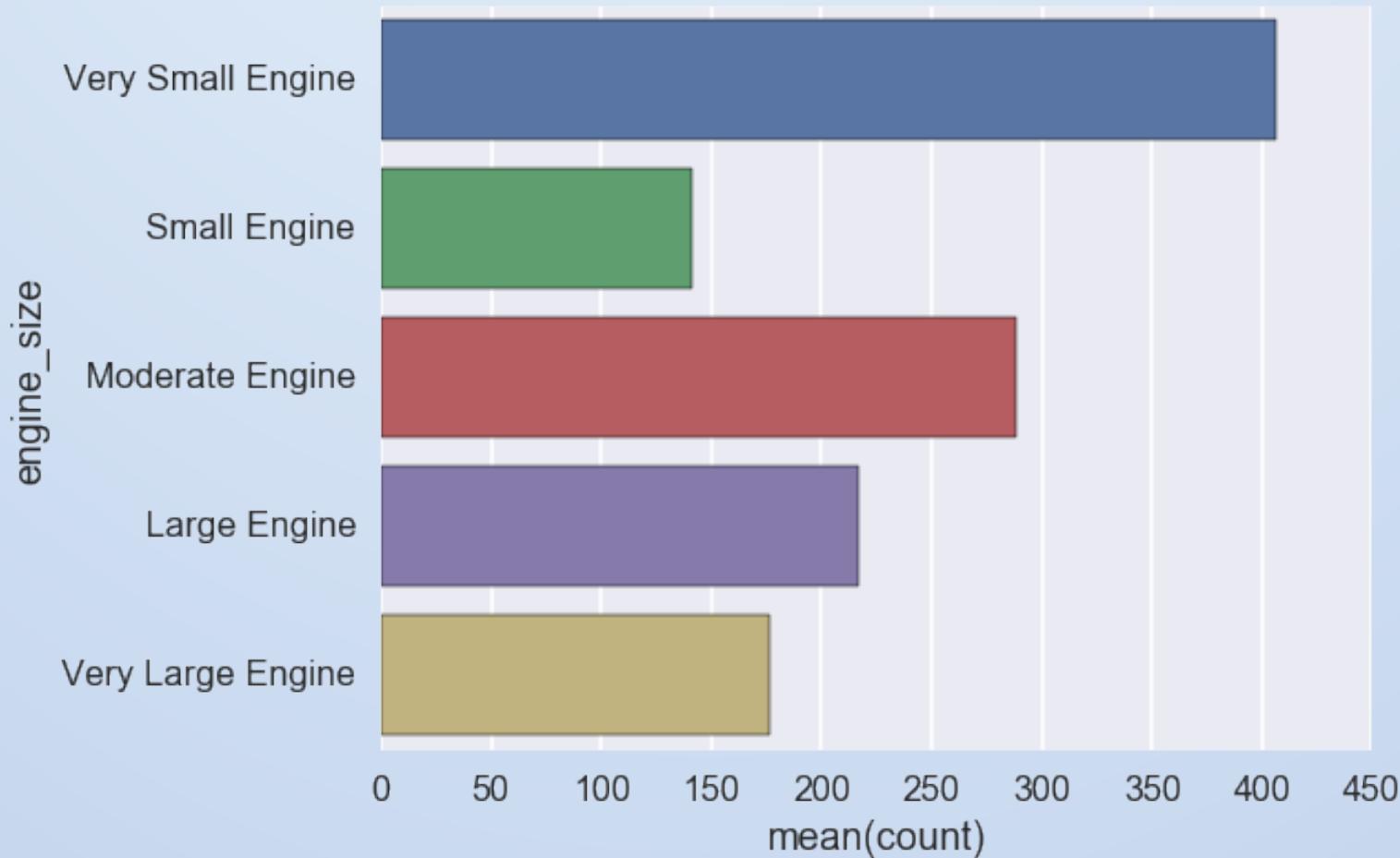
category	count
Small Cars	438
Sport Utility	320
Midsize Cars	201
Large Cars	101
Pickup Trucks	83
Station Wagons	44
Special Purpose	24
Vans & Minivans	11



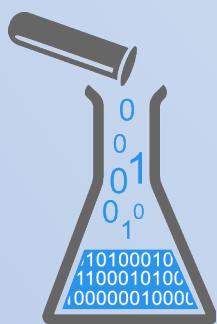
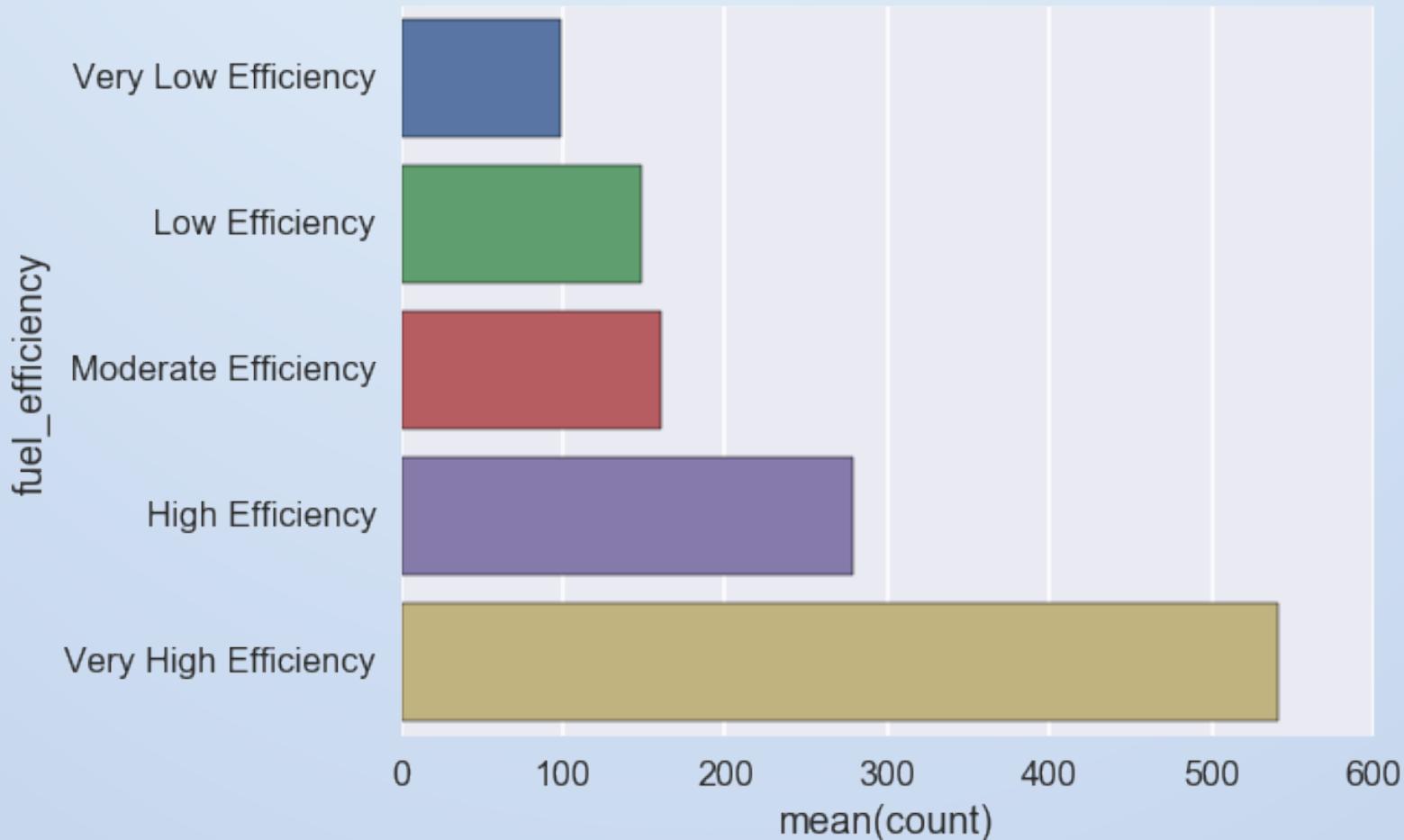
VEHICLE CATEGORY COUNTS (1985)



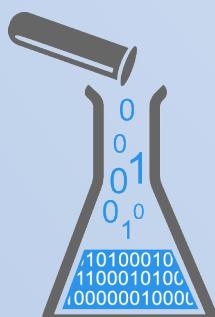
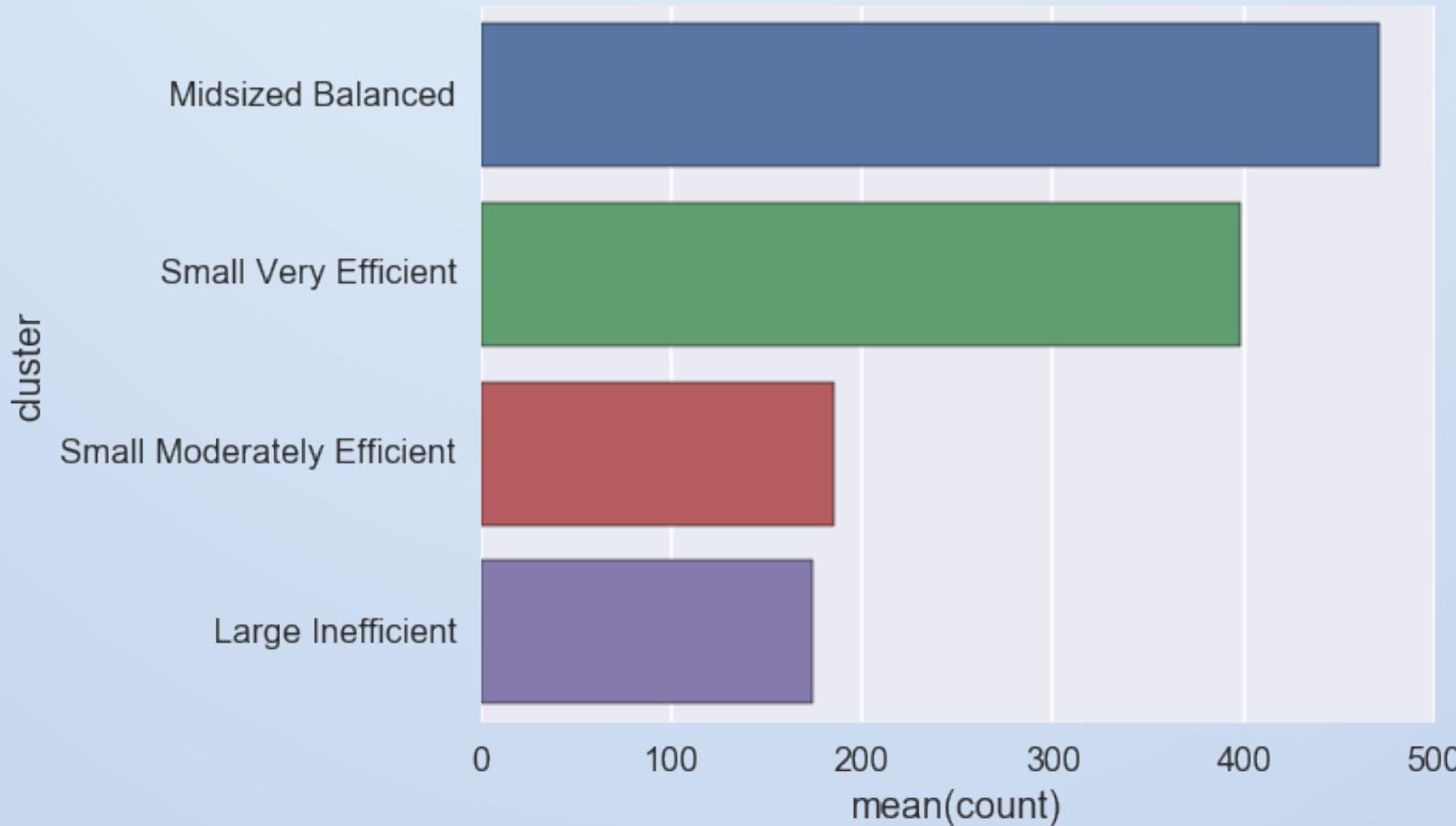
ENGINE SIZE COUNTS (2016)



FUEL EFFICIENCY COUNTS (2016)

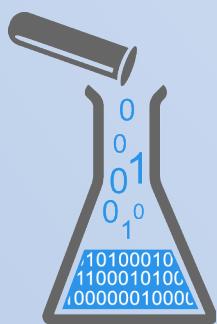
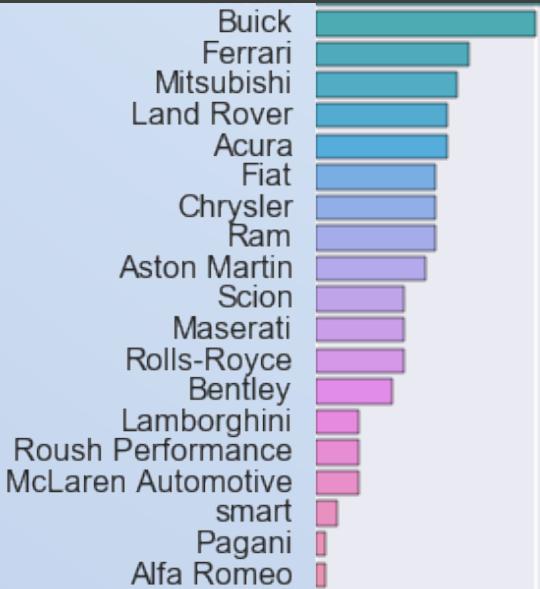


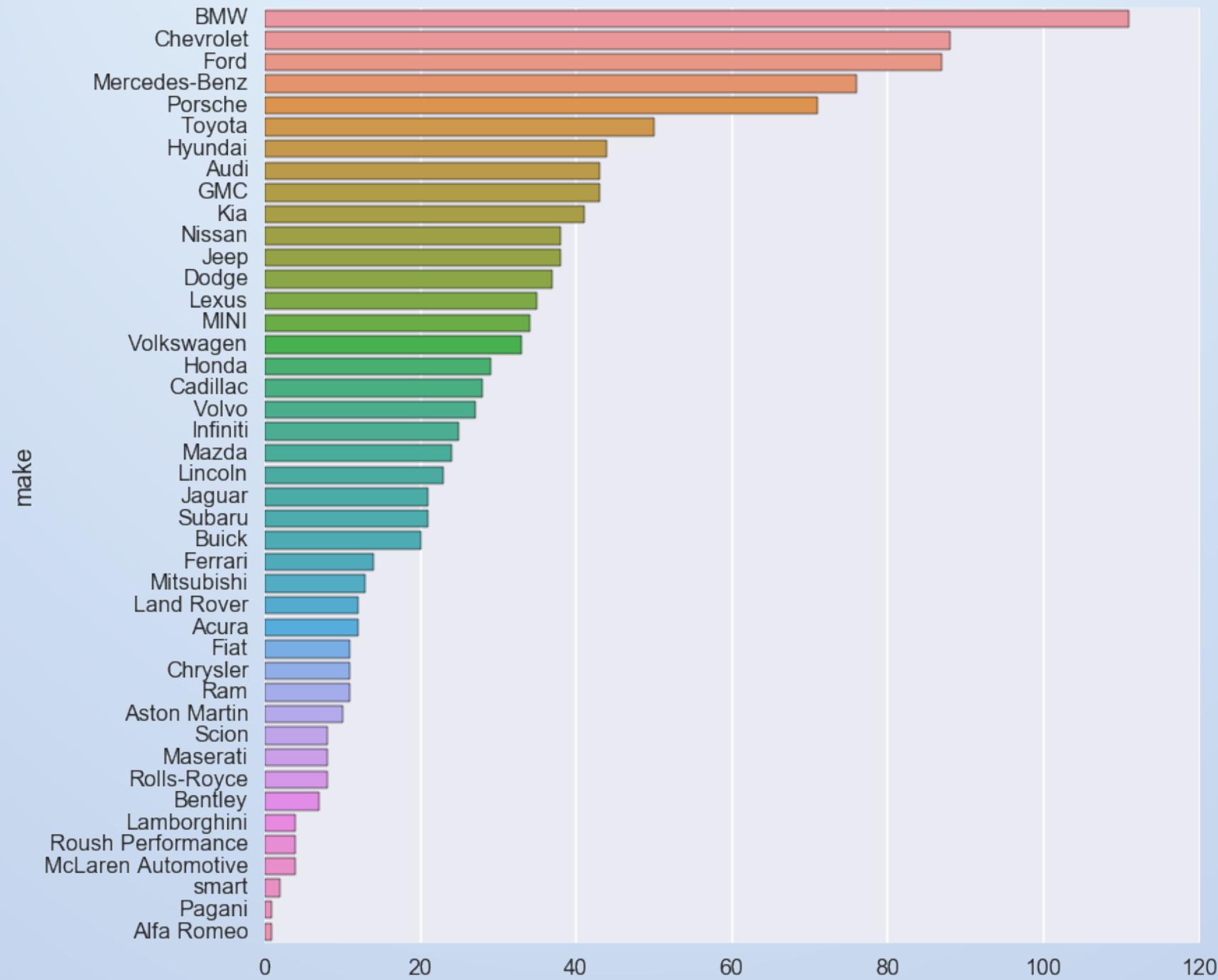
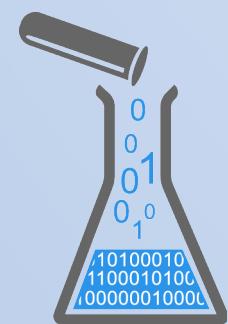
VEHICLE CLUSTER COUNTS (2016)



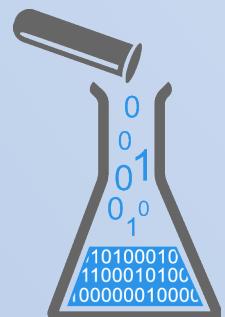
BMW
Chevrolet
Ford
Mercedes-Benz
Porsche
Toyota
Hyundai
Audi
GMC
Kia
Nissan
Jeep
Dodge
Lexus
MINI
Volkswagen
Honda

MANUFACTURER VEHICLE COUNTS (2016)

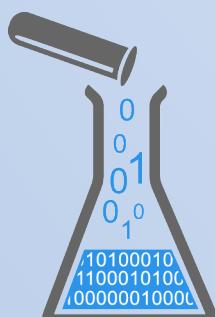
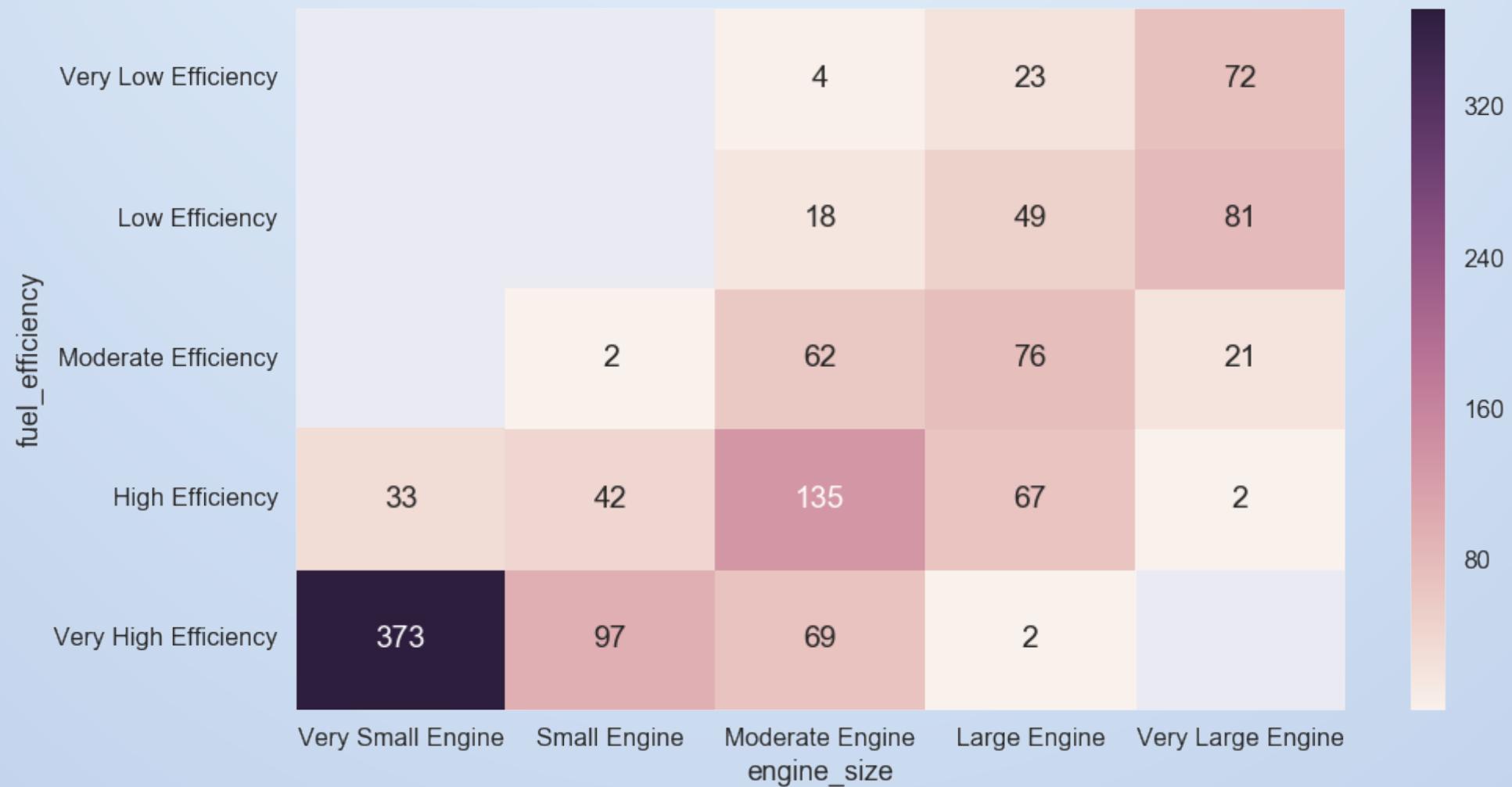




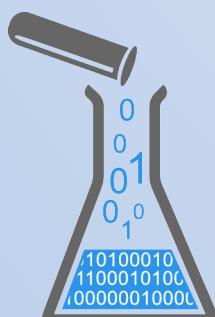
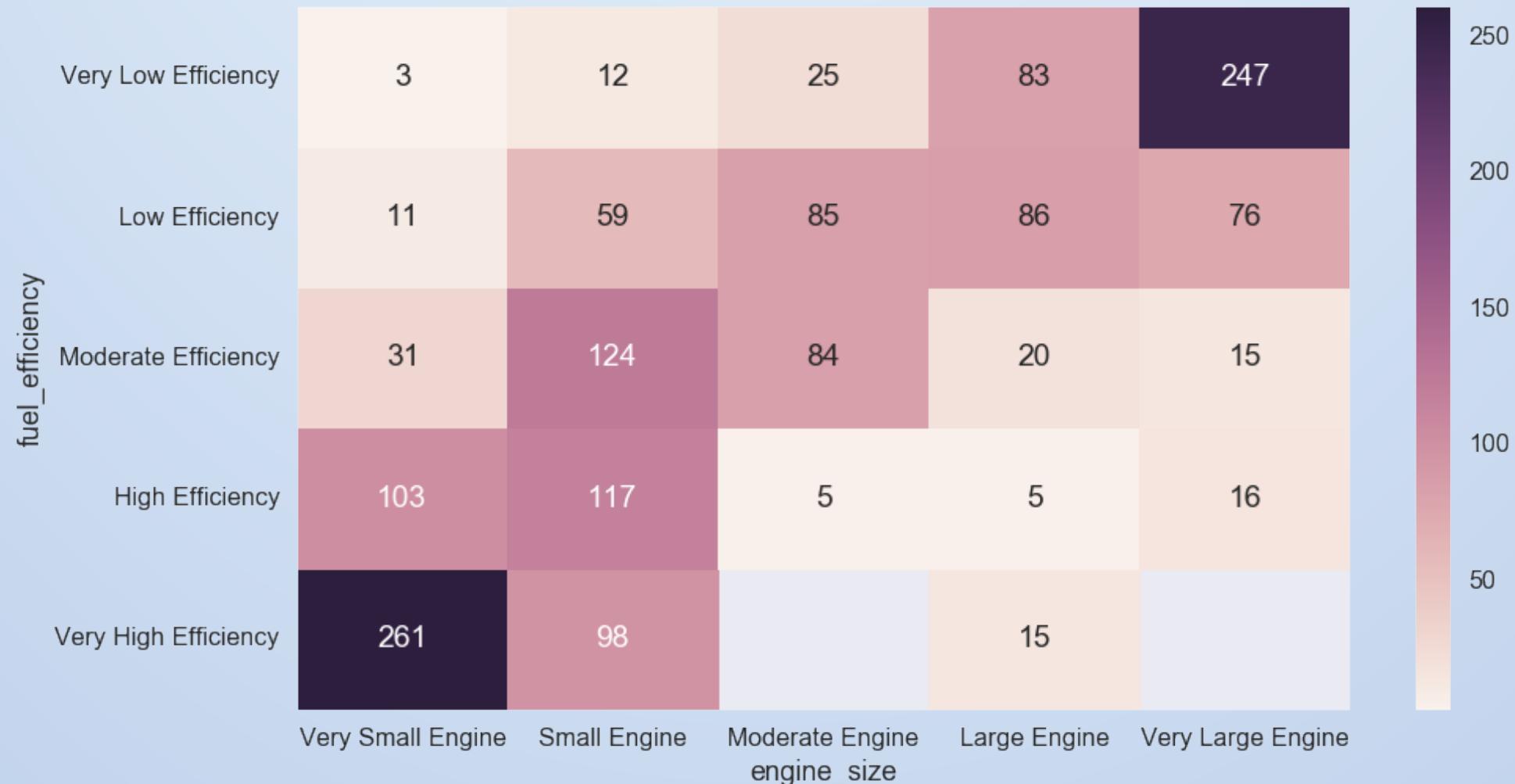
MORE DETAIL



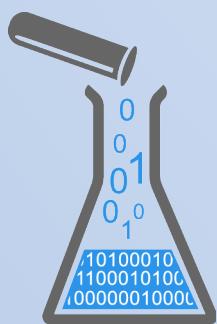
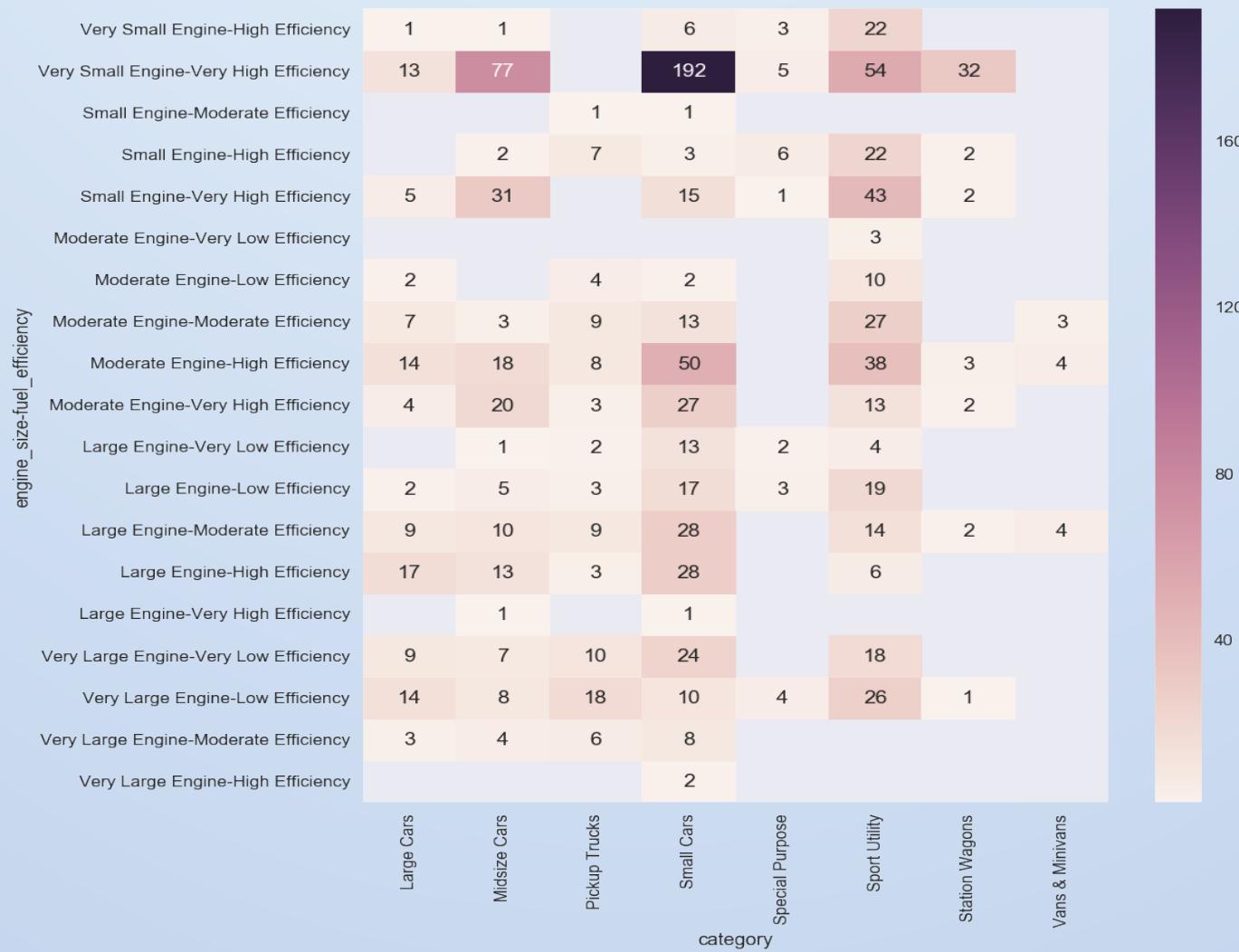
FUEL EFFICIENCY US. ENGINE SIZE (2016)

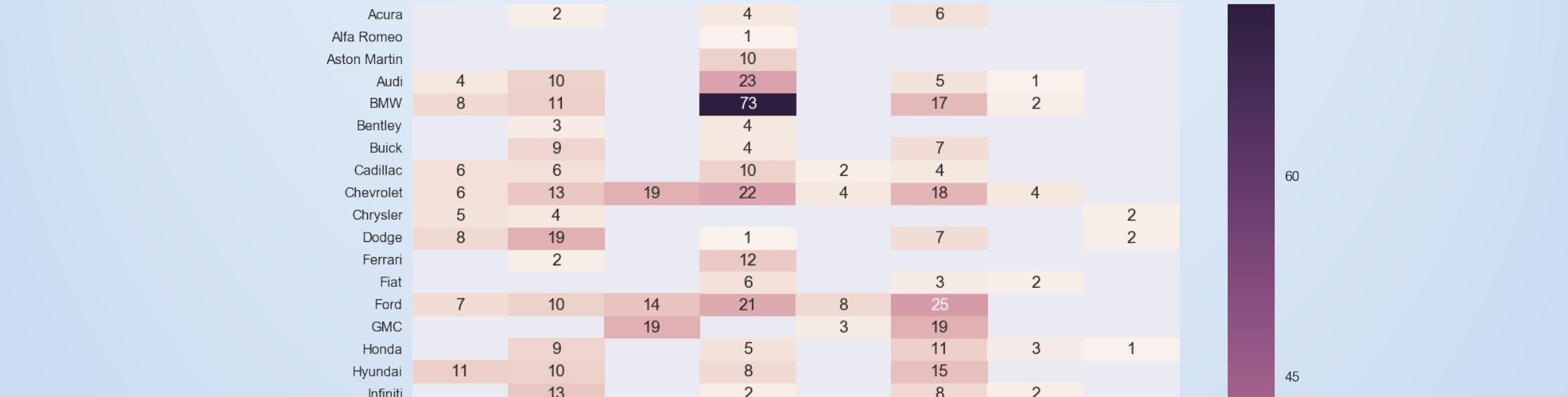


FUEL EFFICIENCY US. ENGINE SIZE (1985)

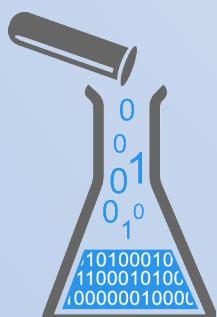
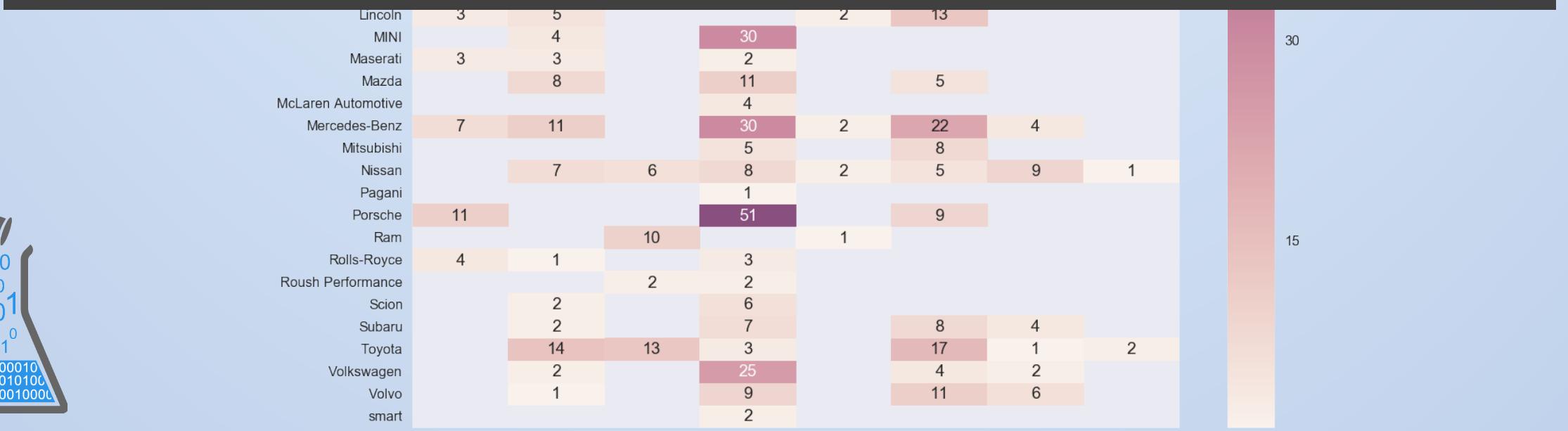


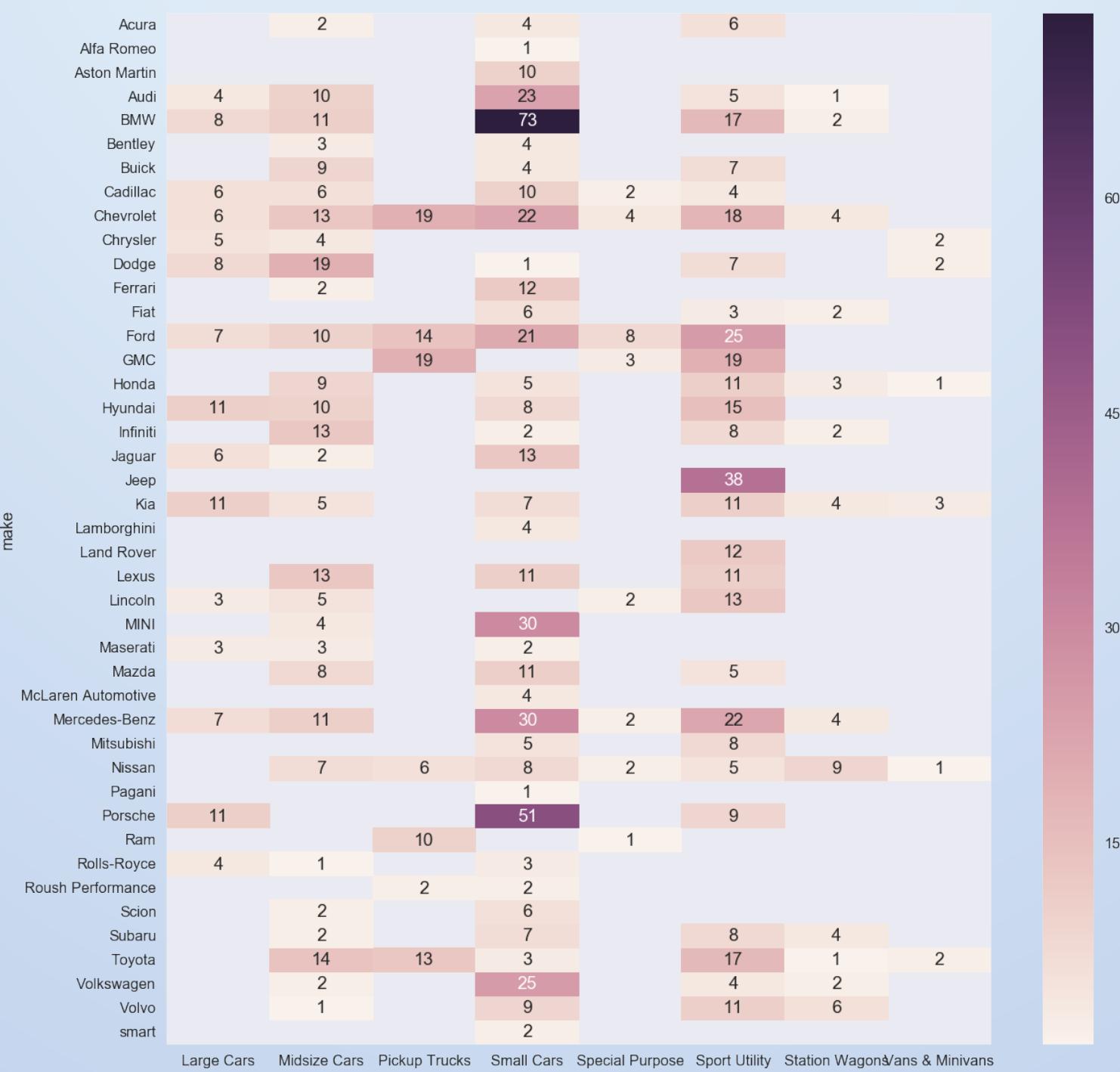
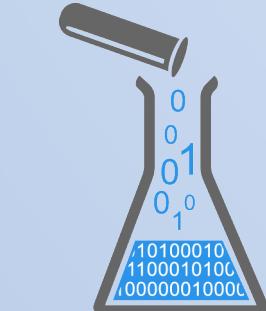
ENGINE SIZE & EFFICIENCY US. CATEGORY



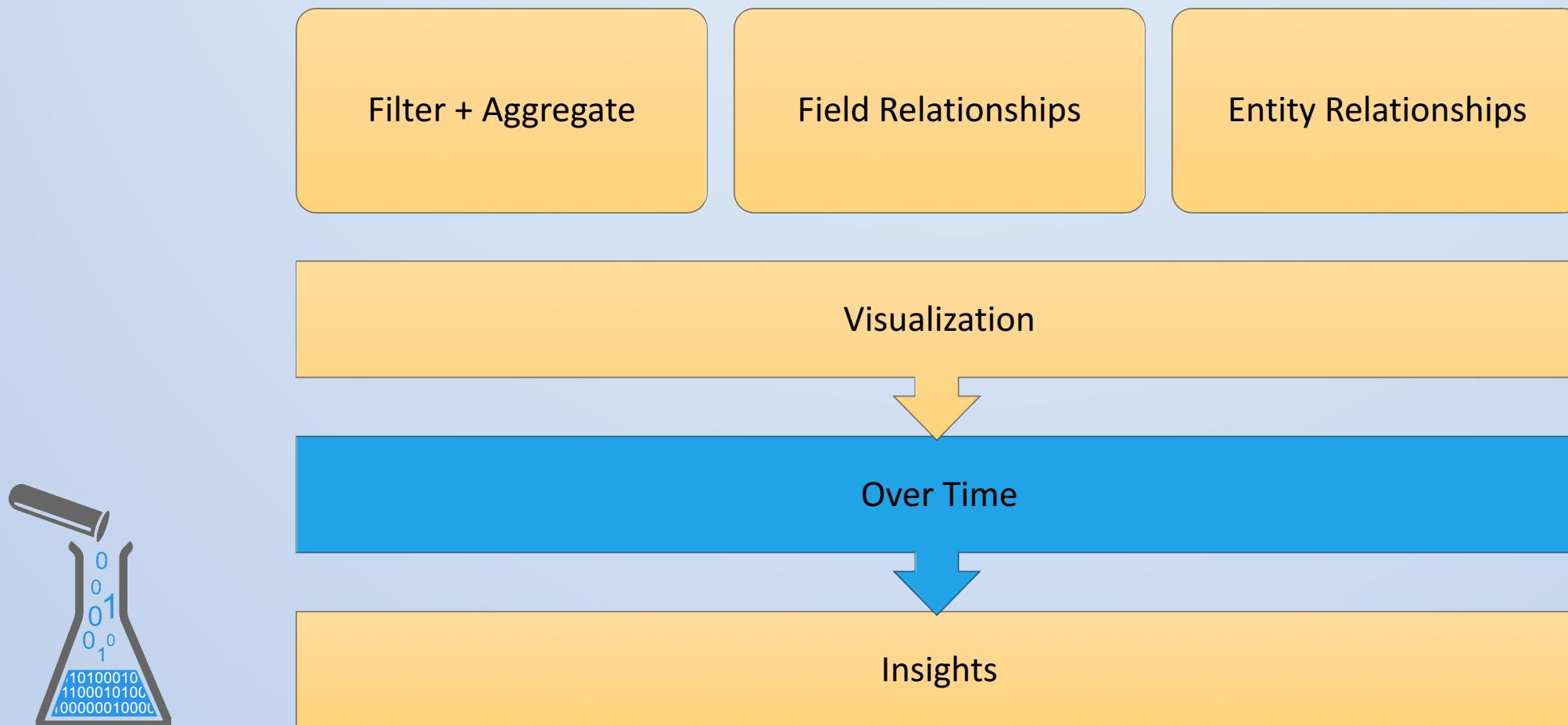


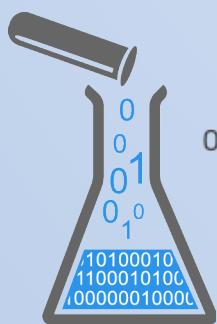
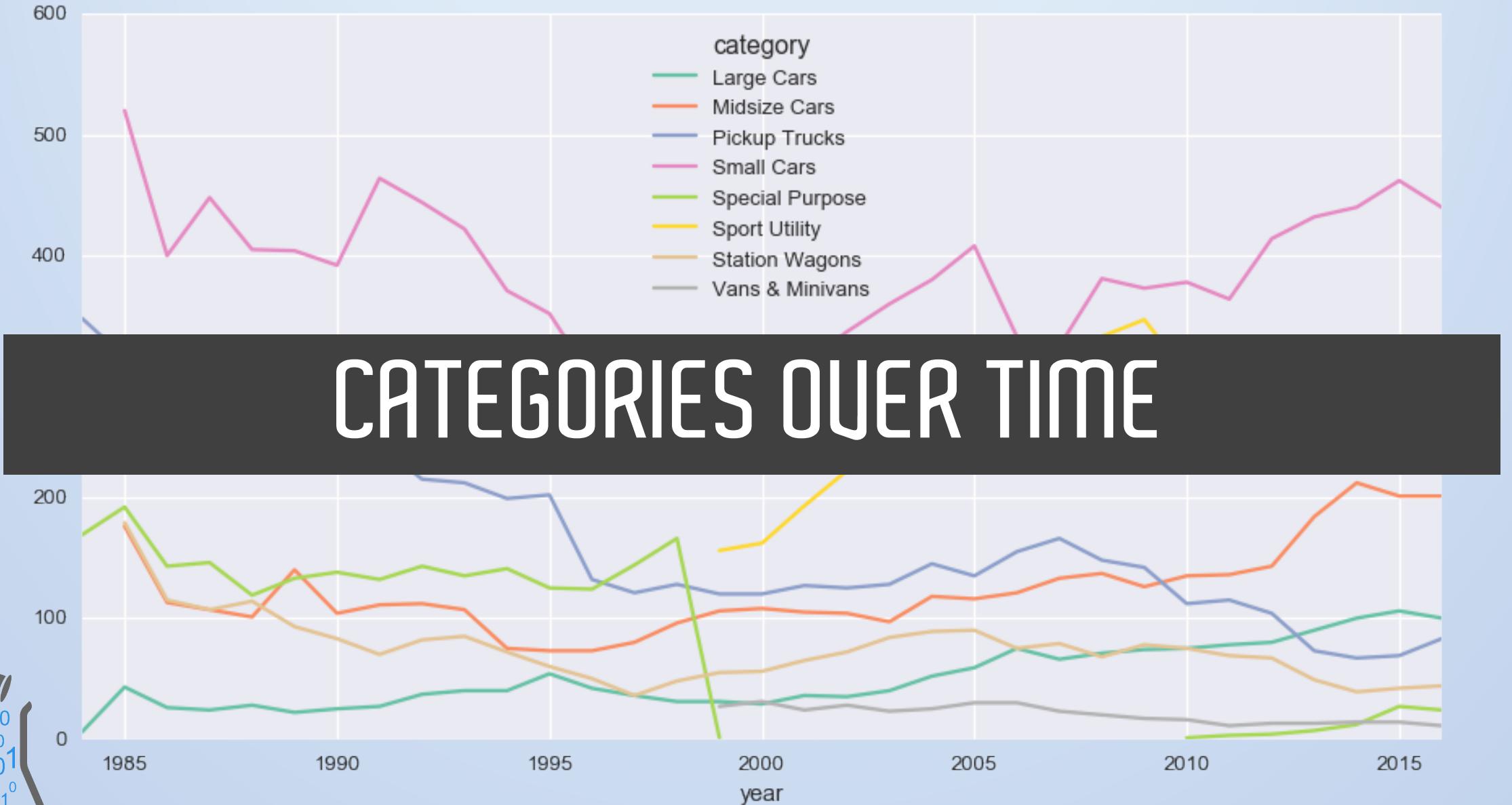
PIUOT COUNTS BY MAKE & CATEGORY

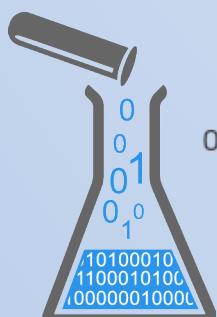
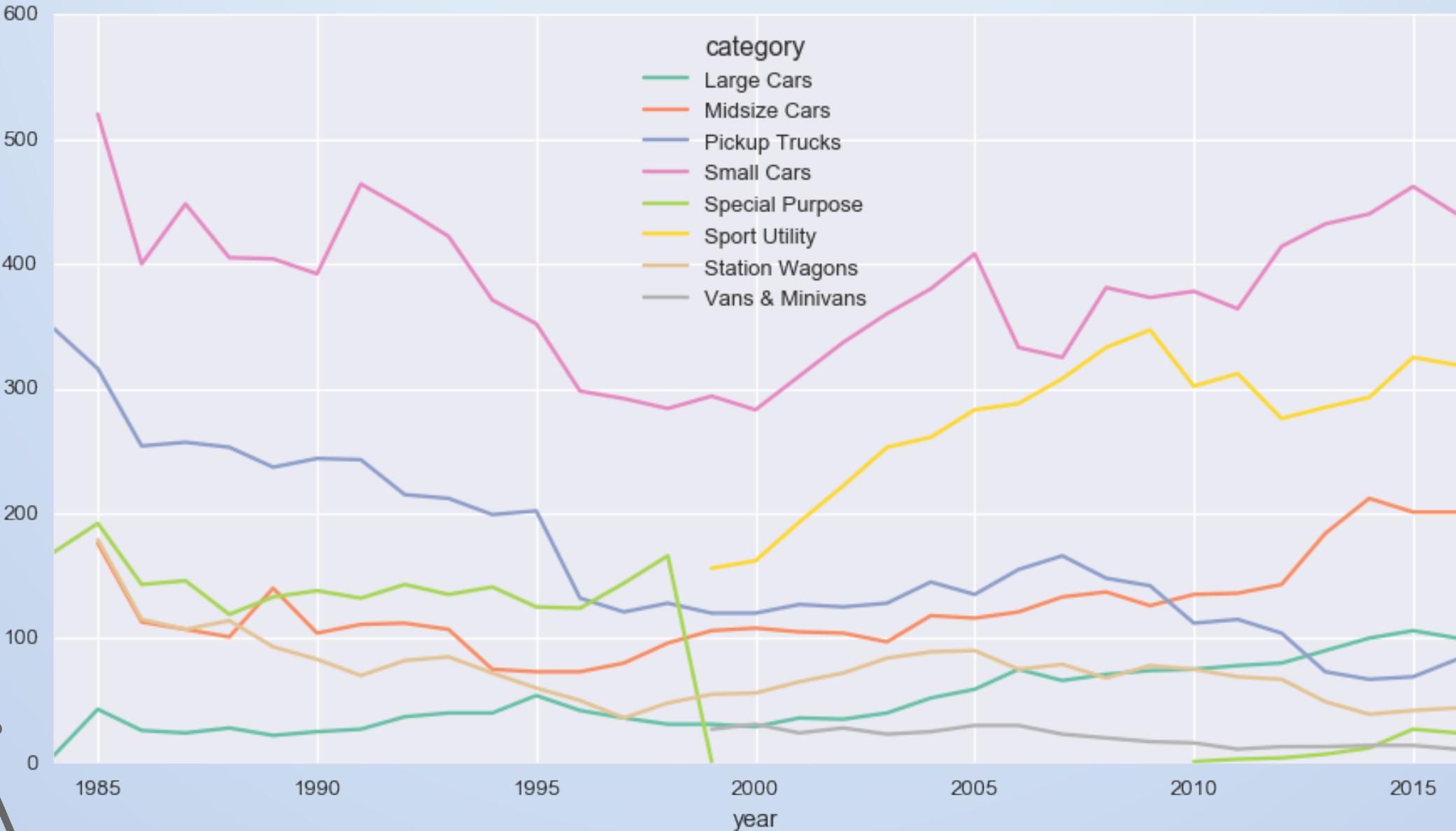


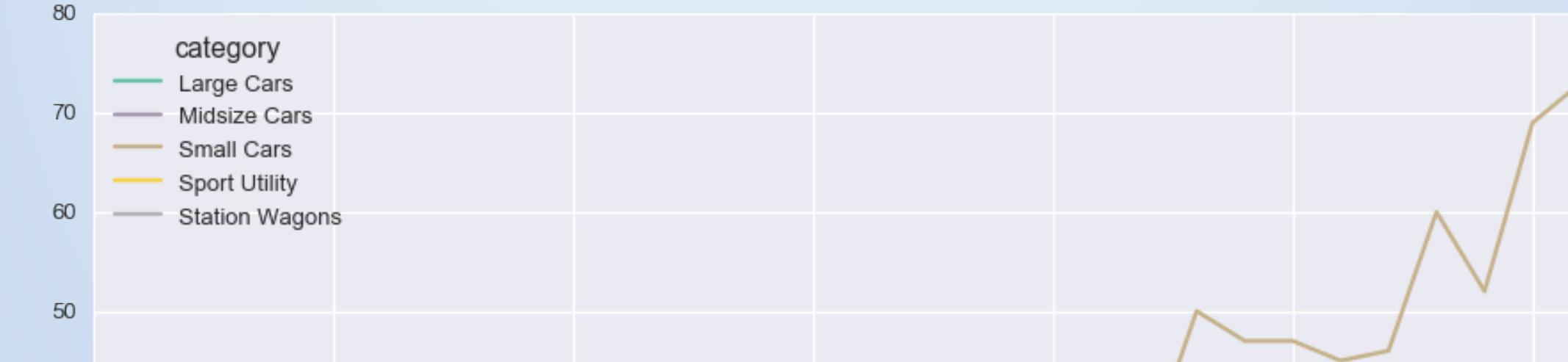


CHANGES OVER TIME

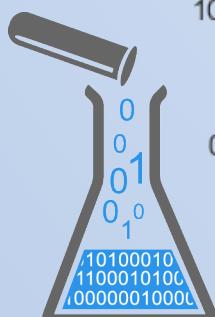
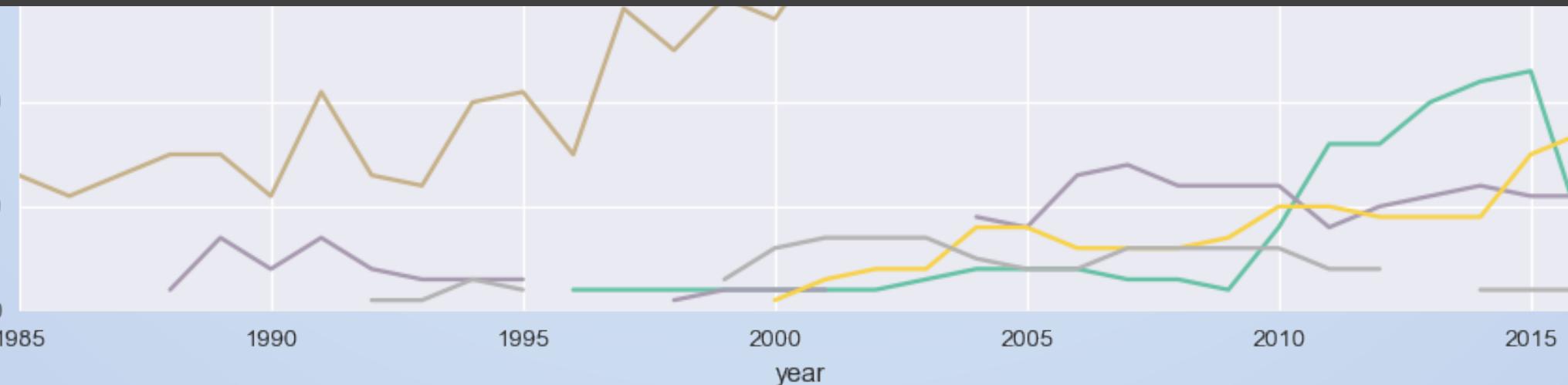


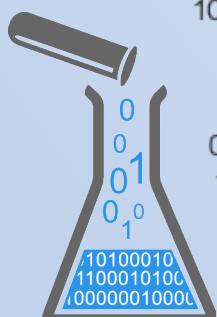
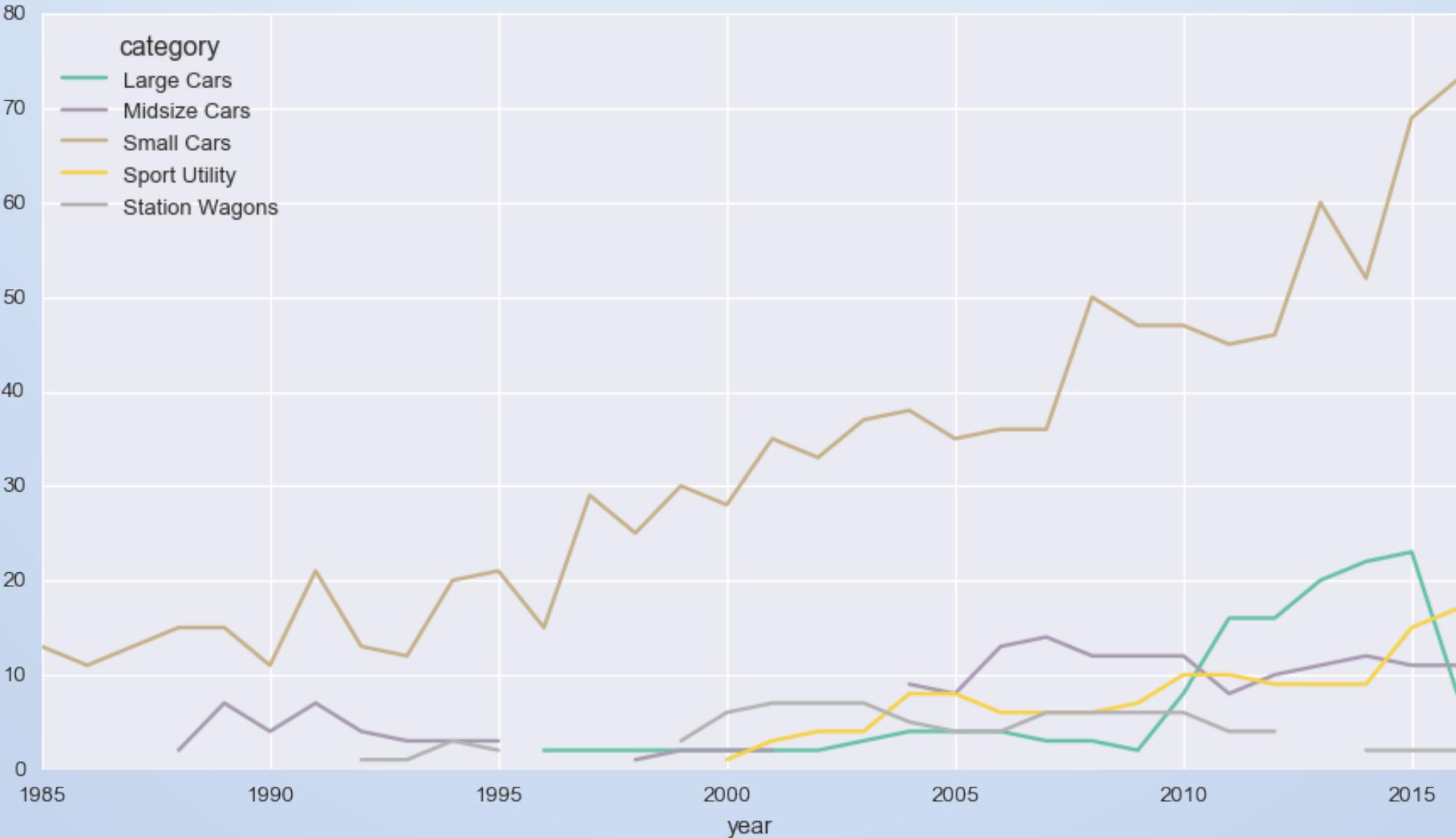


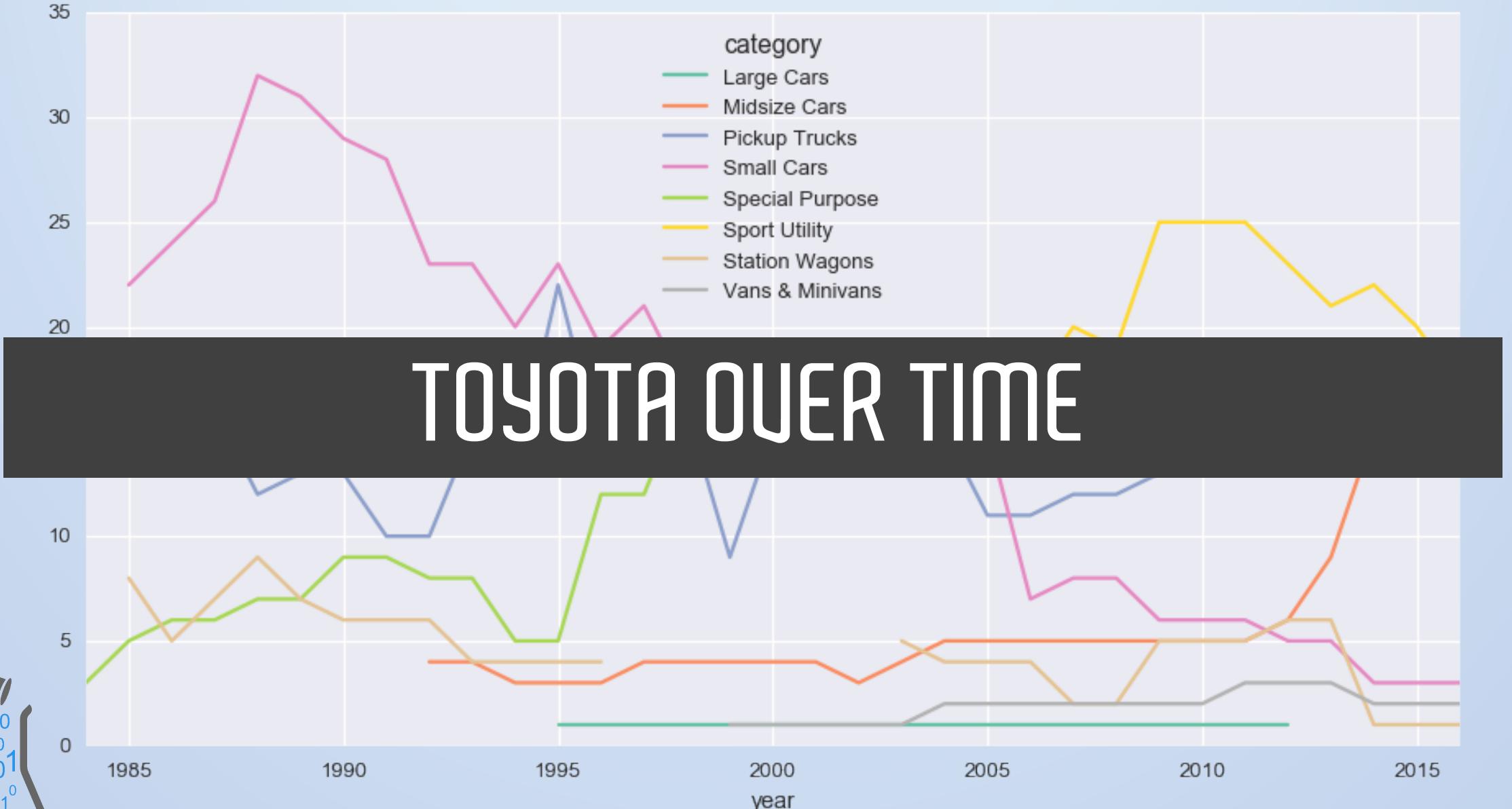




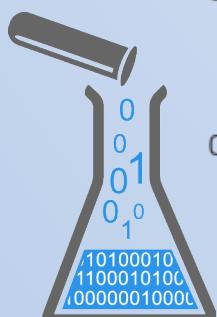
BMW OVER TIME

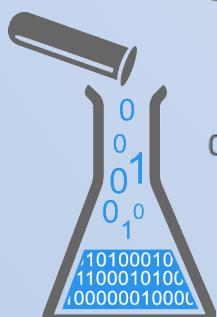
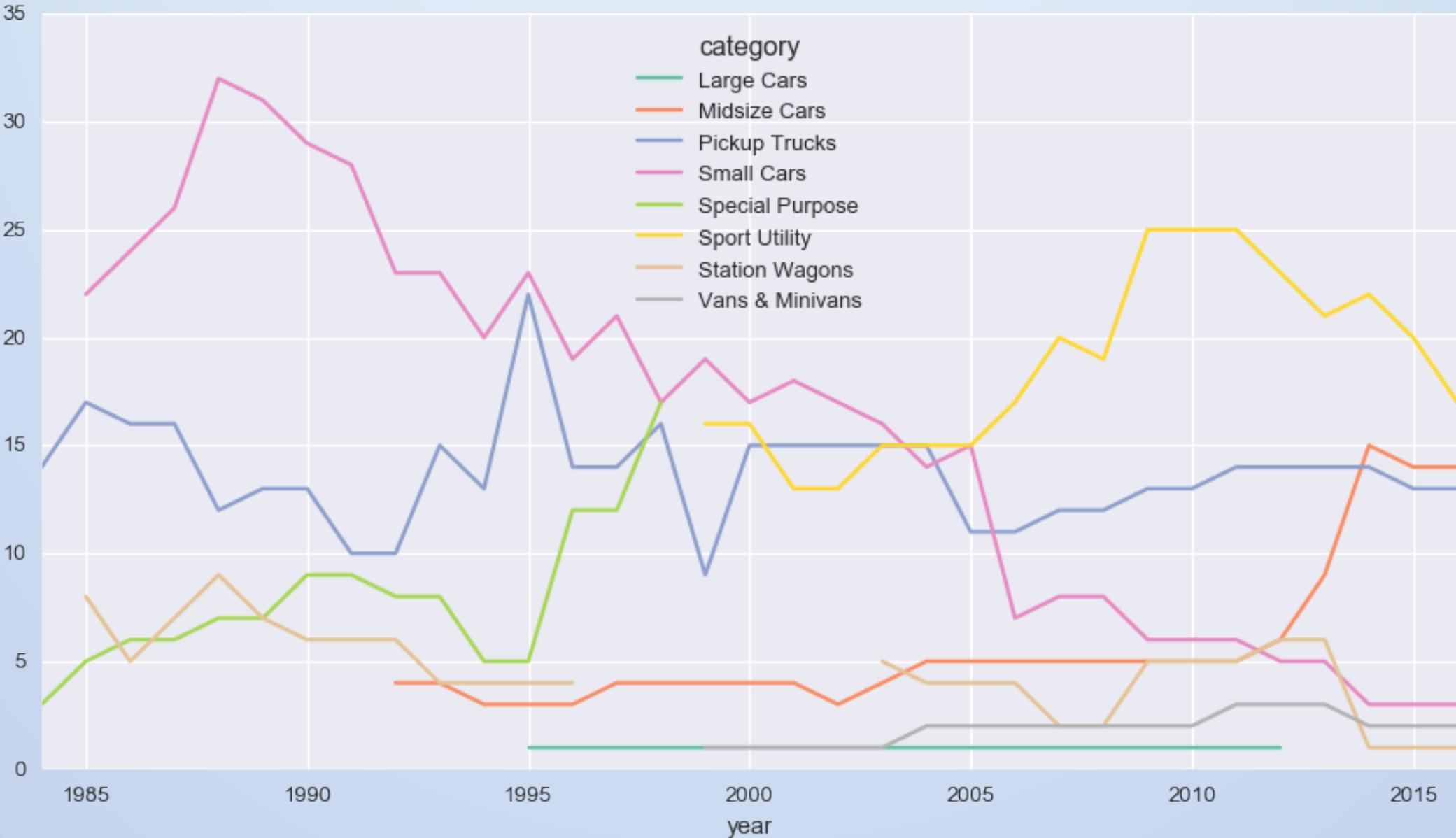






TOYOTA OVER TIME





EXPLORE PHASE

Filter + Aggregate

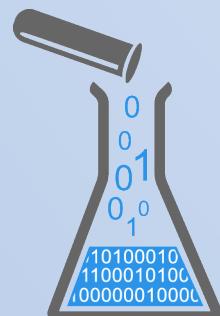
Field Relationships

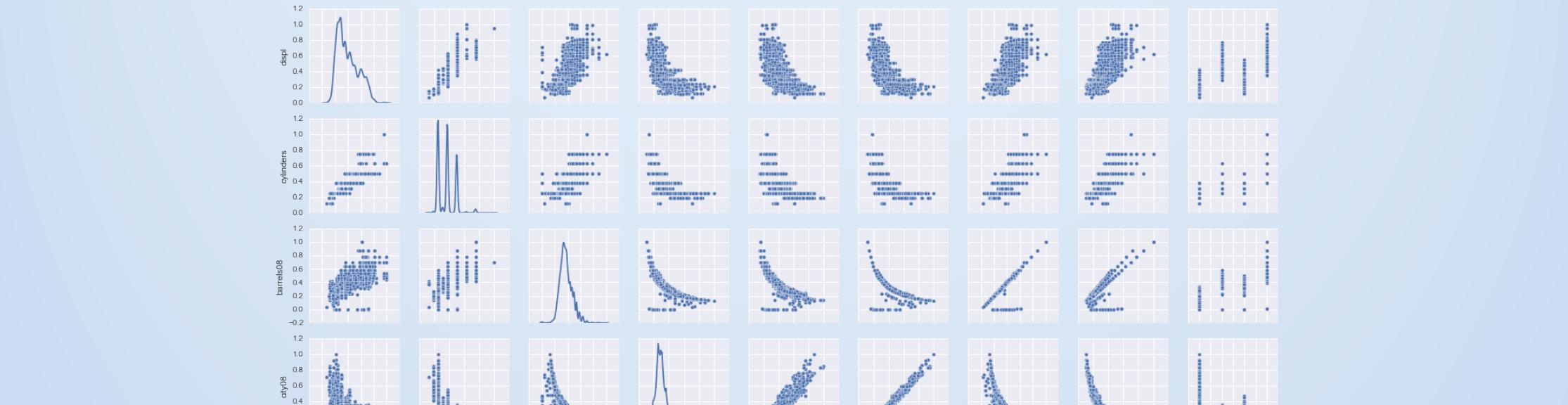
Entity Relationships

Visualization

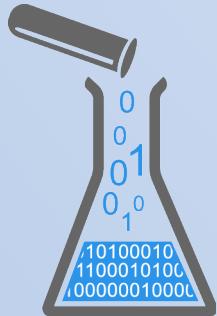
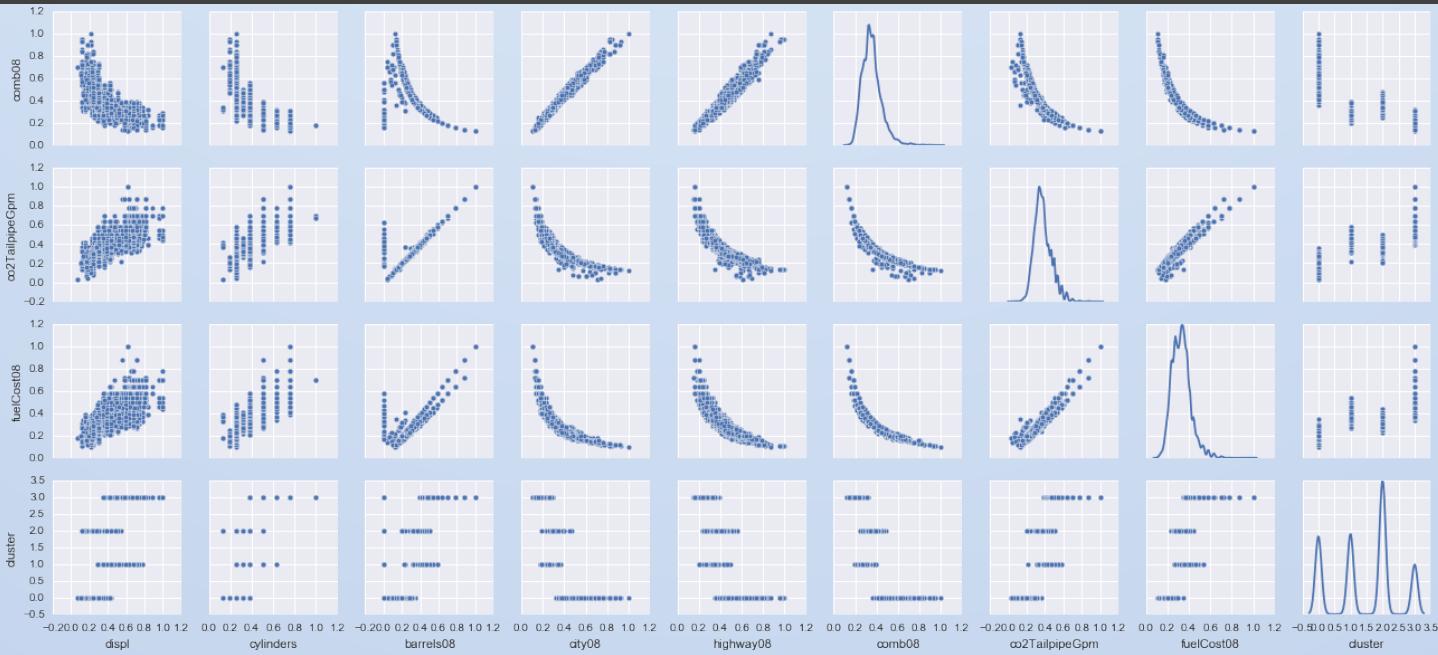
Over Time

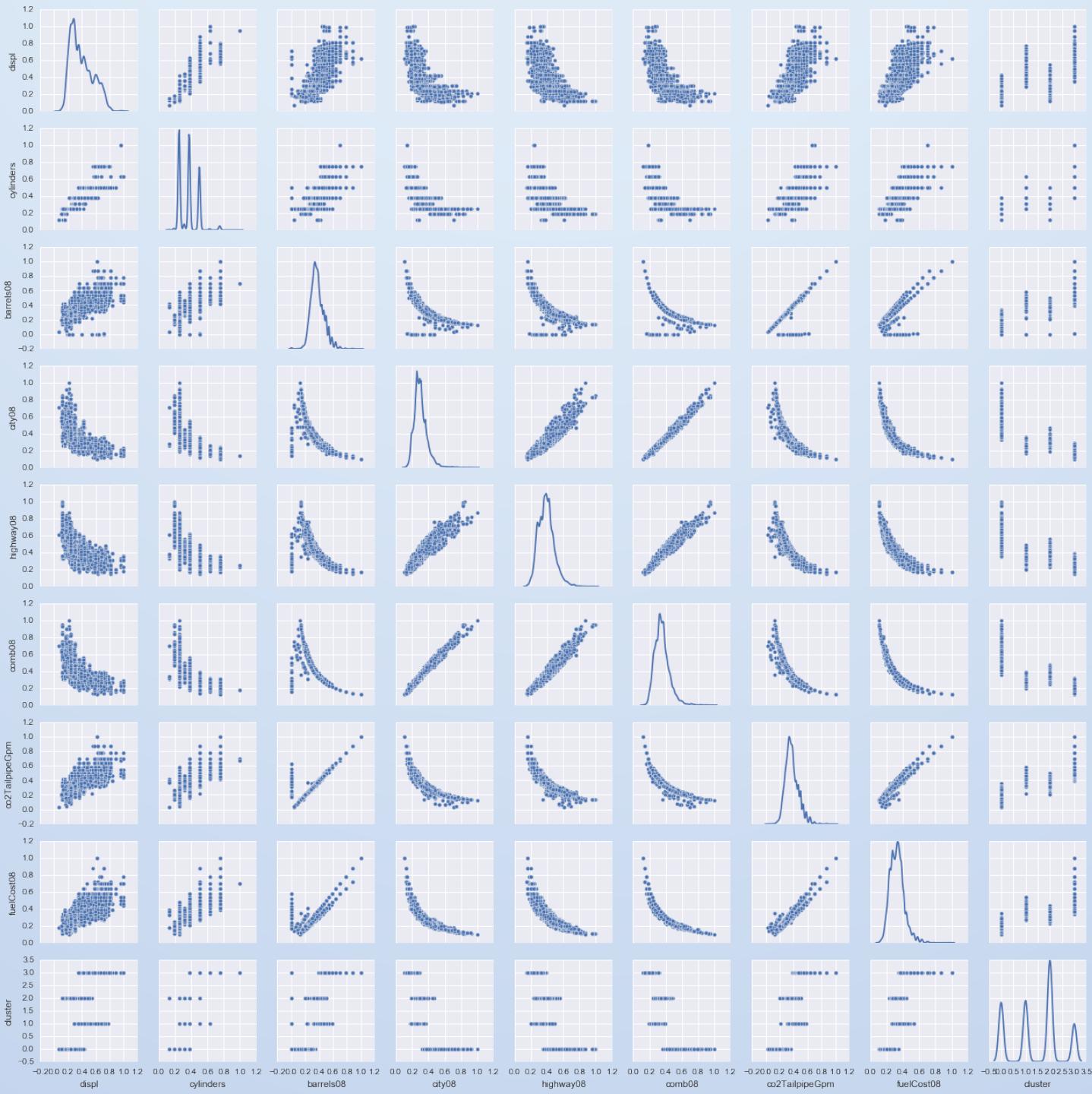
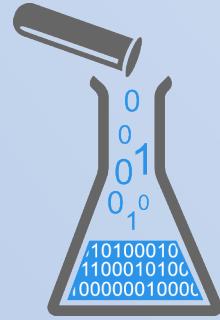
Insights





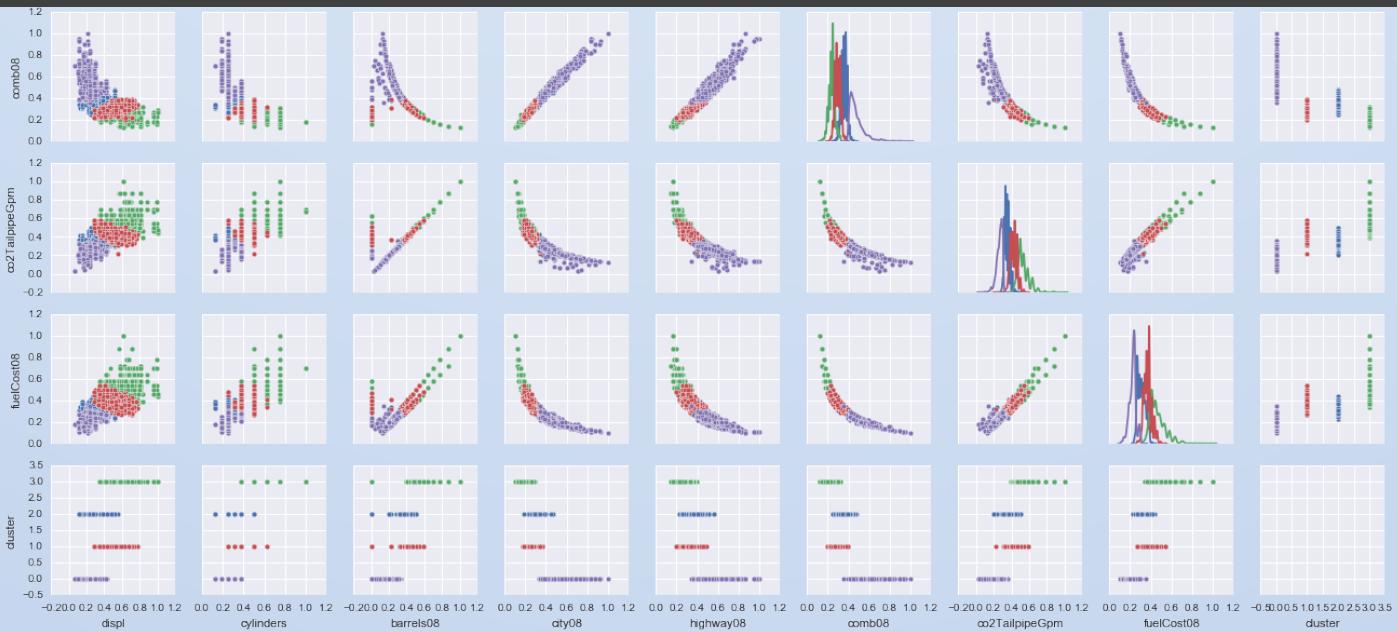
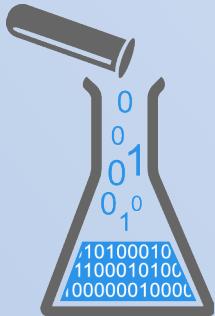
SCATTER MATRICES & PLOTS

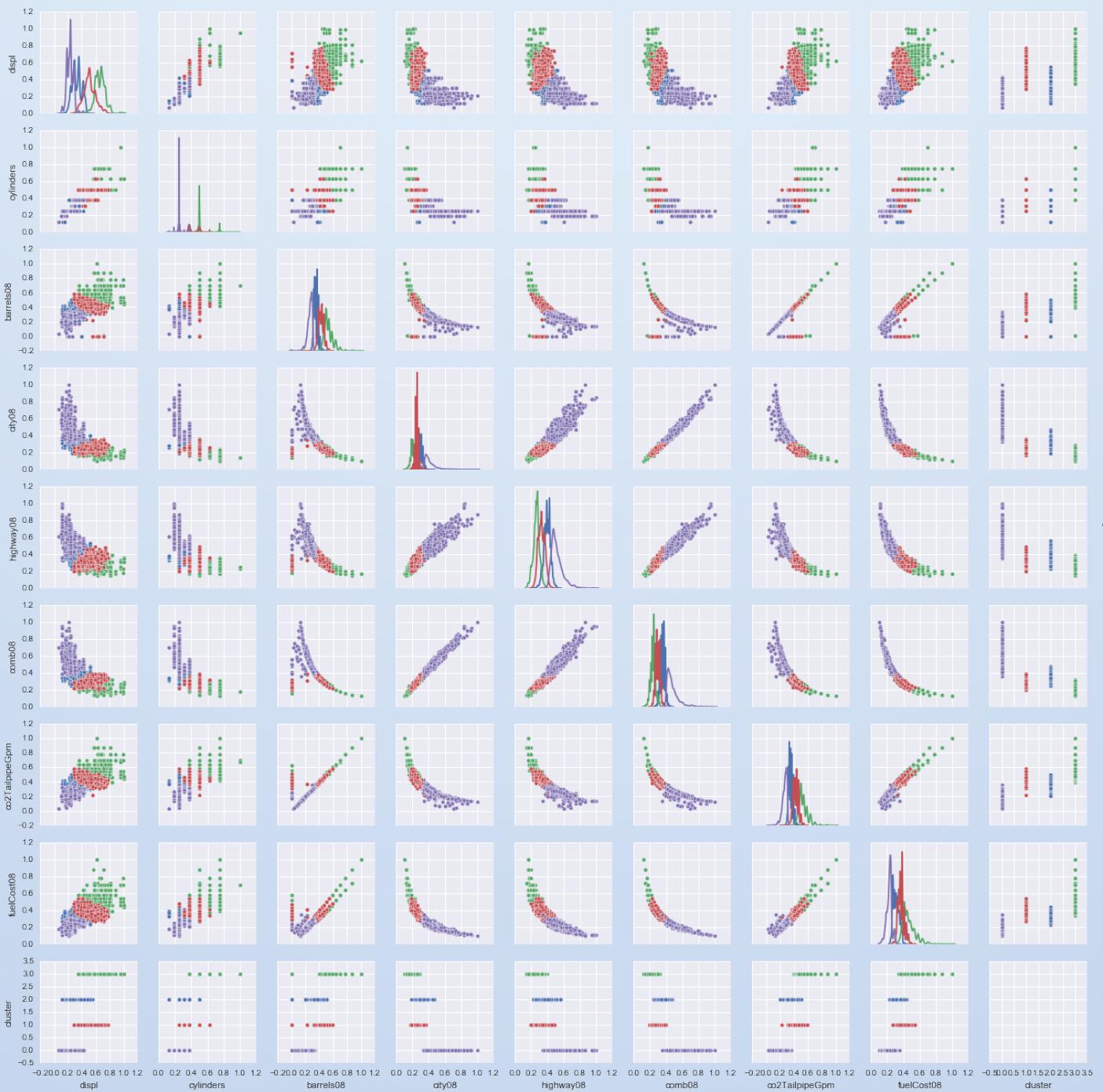
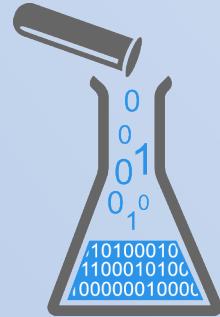


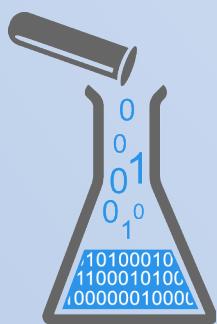


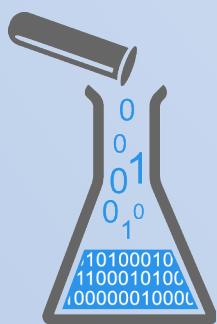
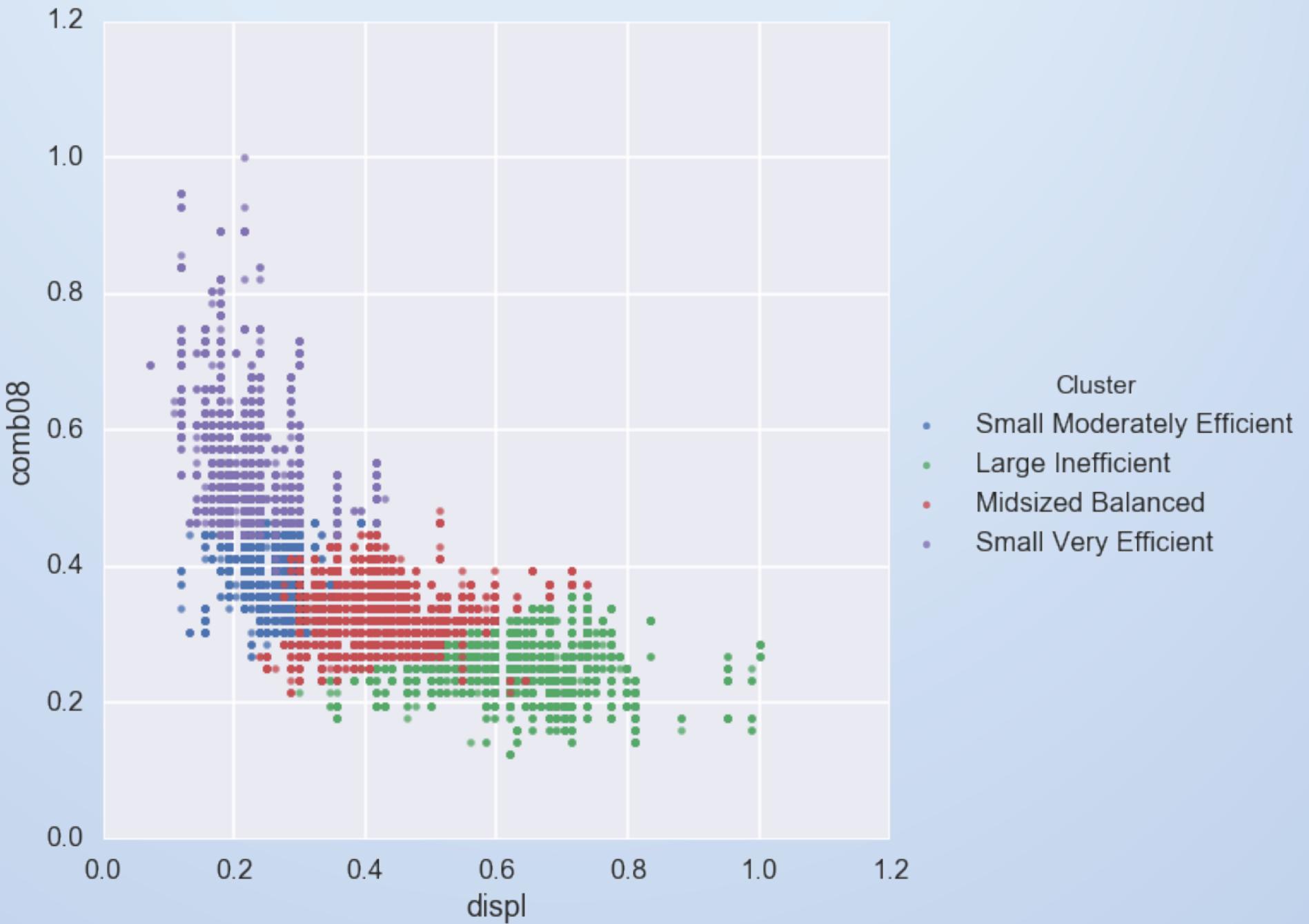


SCATTER MATRIX WITH CATEGORIES

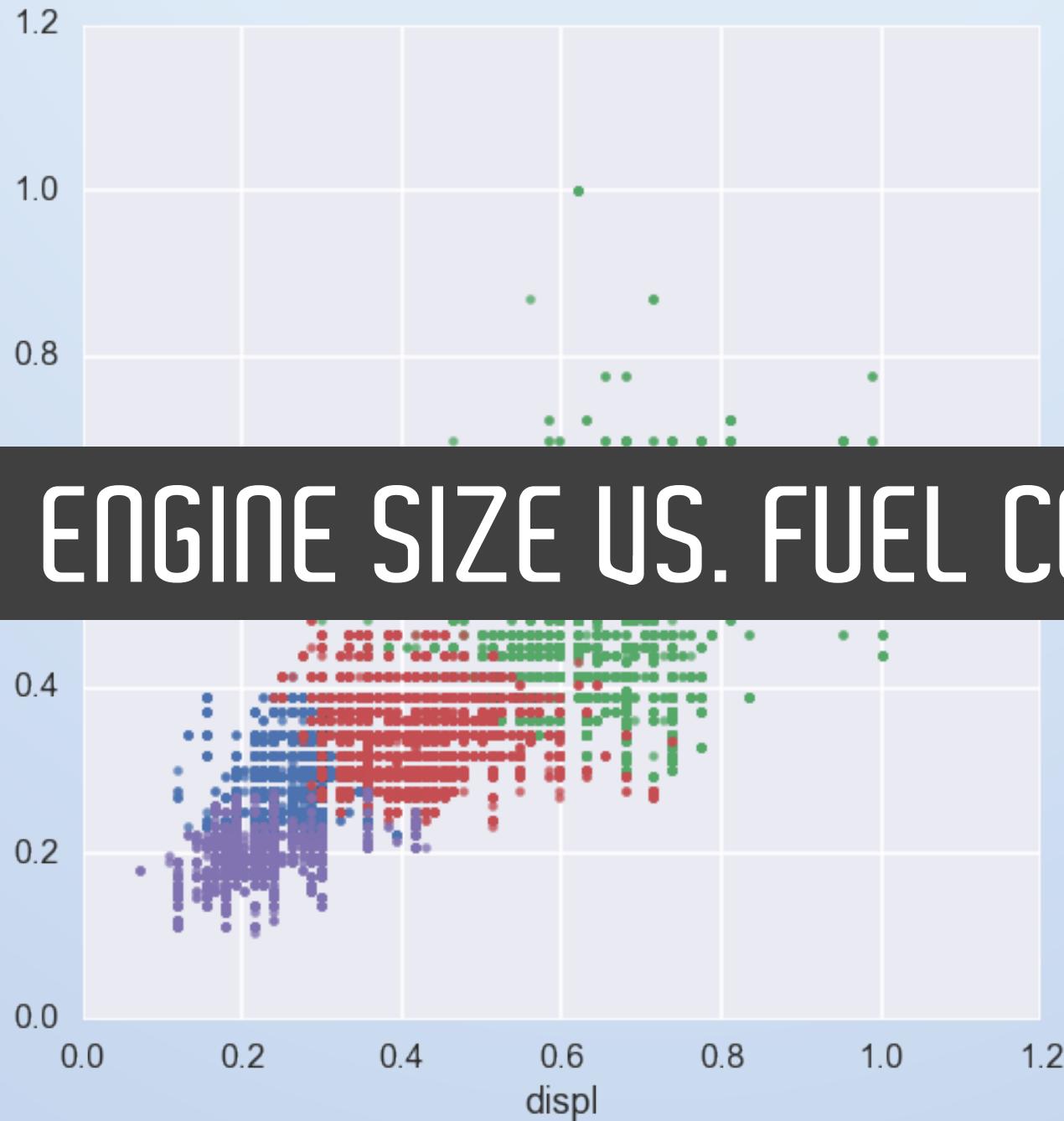
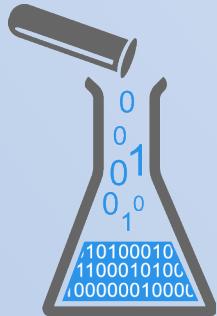


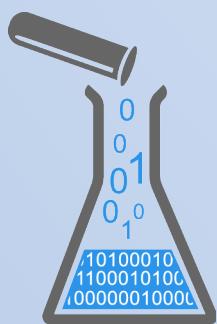
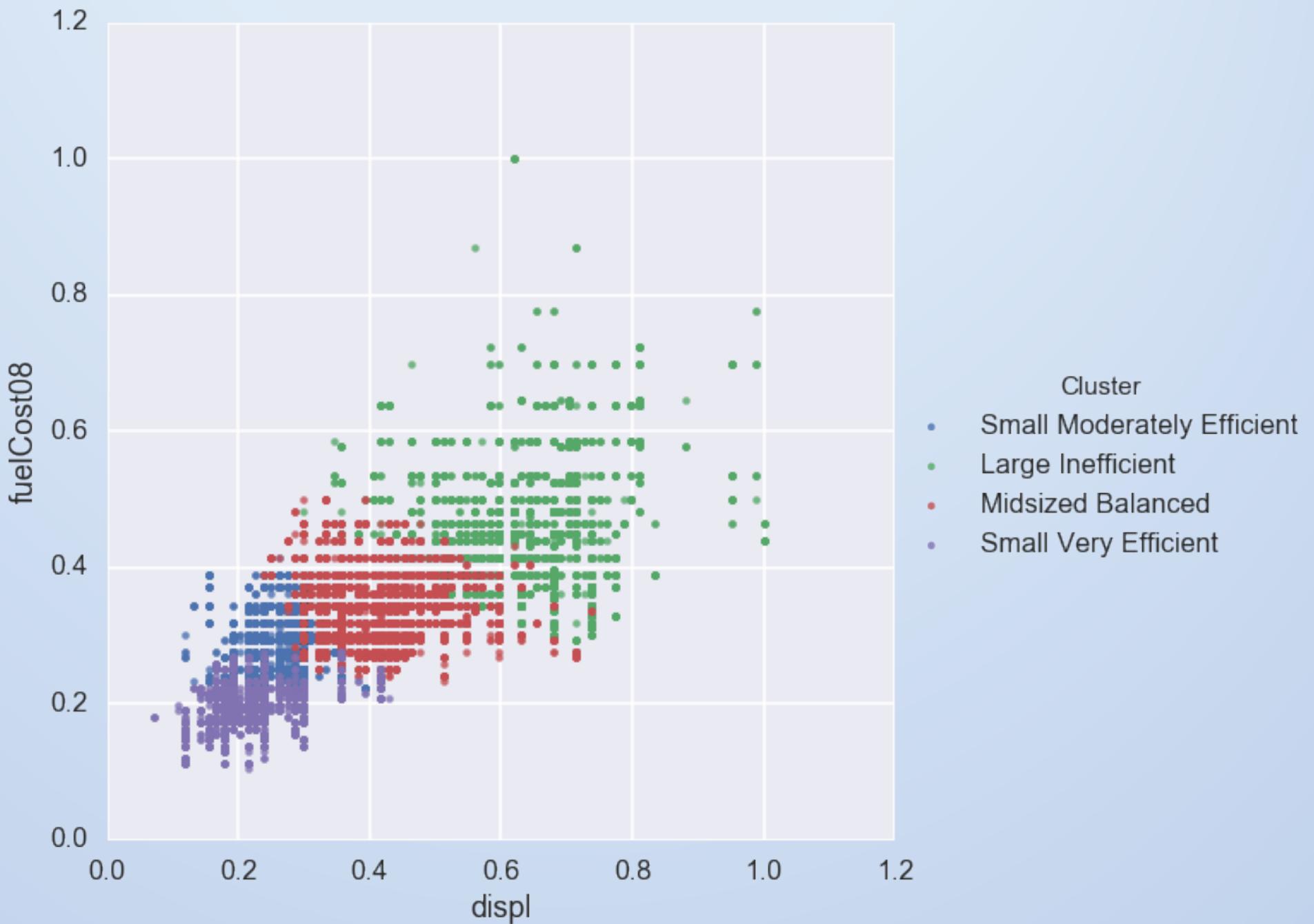






ENGINE SIZE US. FUEL COST





10100010
1100010100
00000010000

EXPLORE PHASE

Filter + Aggregate

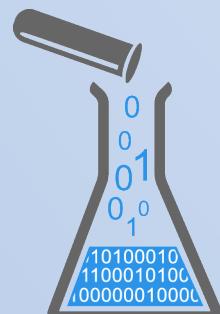
Field Relationships

Entity Relationships

Visualization

Over Time

Insights

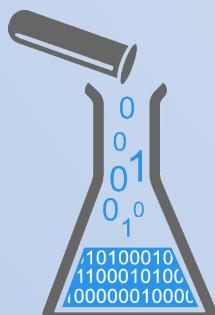


GRAPH ANALYSIS

Relationships between entities.

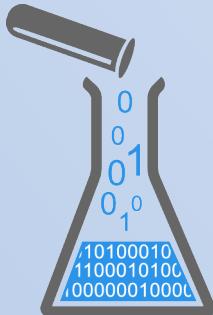
Attributes entities have in common.

Changes in relationships over time.



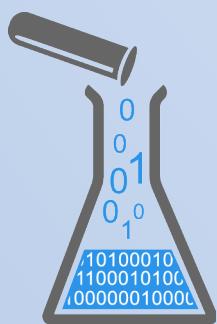
RELATIONAL TO GRAPH TRANSFORMATION

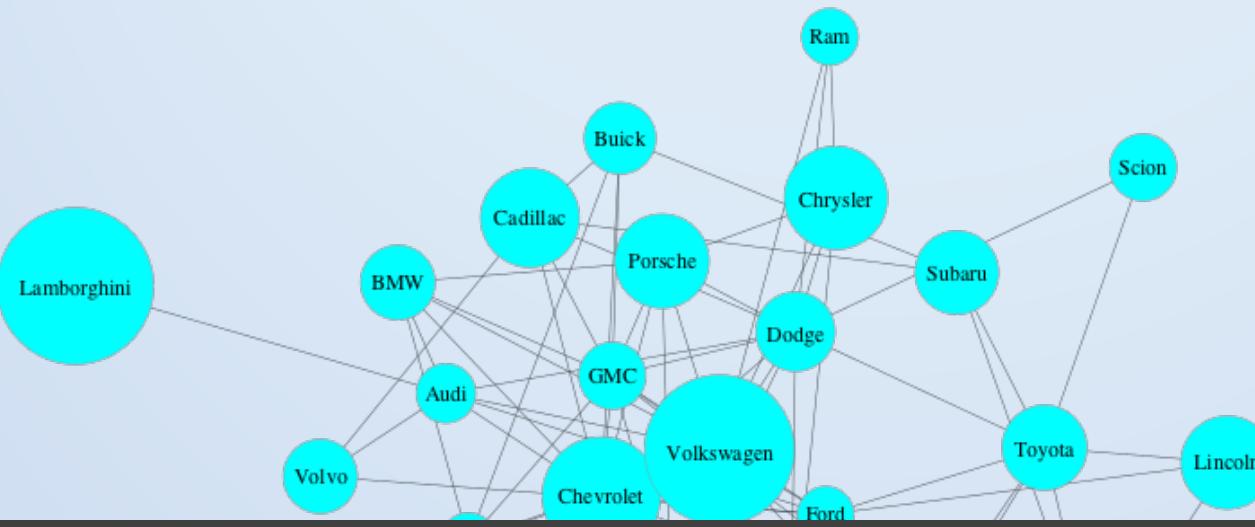
make	model	year	displ	cylinders	trany	drive	VClass	fuelType	barrels08	city08	highway08	comb08	co2TailpipeGpm	fuelCost08
AM General	DJ Po Vehicle 2WD	1984	2.5	4	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	19.388824	18	17	17	522.764706	2000
AM General	FJ8c Post Office	1984	4.2	6	Automatic 3-spd	2-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
AM General	Post Office DJ5 2WD	1985	2.5	4	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	20.600625	16	17	16	555.4375	2100
AM General	Post Office DJ8 2WD	1985	4.2	6	Automatic 3-spd	Rear-Wheel Drive	Special Purpose Vehicle 2WD	Regular	25.354615	13	13	13	683.615385	2600
ASC Incorporated	GNX	1987	3.8	6	Automatic 4-spd	Rear-Wheel Drive	Midsize Cars	Premium	20.600625	14	21	16	555.4375	2550
Acura	2.2CL/3.0CL	1997	2.2	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	26	22	403.954545	1550
Acura	2.2CL/3.0CL	1997	2.2	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	22	28	24	370.291667	1400
Acura	2.2CL/3.0CL	1997	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	18	26	20	444.35	1700
Acura	2.3CL/3.0CL	1998	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	19	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1998	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1998	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.3CL/3.0CL	1999	2.3	4	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	14.982273	20	27	22	403.954545	1550
Acura	2.3CL/3.0CL	1999	2.3	4	Manual 5-spd	Front-Wheel Drive	Subcompact Cars	Regular	13.73375	21	29	24	370.291667	1400
Acura	2.3CL/3.0CL	1999	3	6	Automatic 4-spd	Front-Wheel Drive	Subcompact Cars	Regular	16.4805	17	26	20	444.35	1700
Acura	2.5TL	1995	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1996	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1997	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	18	23	20	444.35	2050
Acura	2.5TL/3.2TL	1997	3.2	6	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	17.347895	17	22	19	467.736842	2150
Acura	2.5TL/3.2TL	1998	2.5	5	Automatic 4-spd	Front-Wheel Drive	Compact Cars	Premium	16.4805	17	23	20	444.35	2050



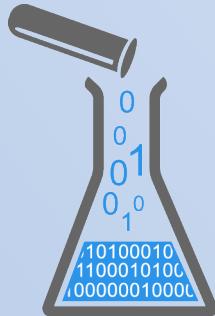
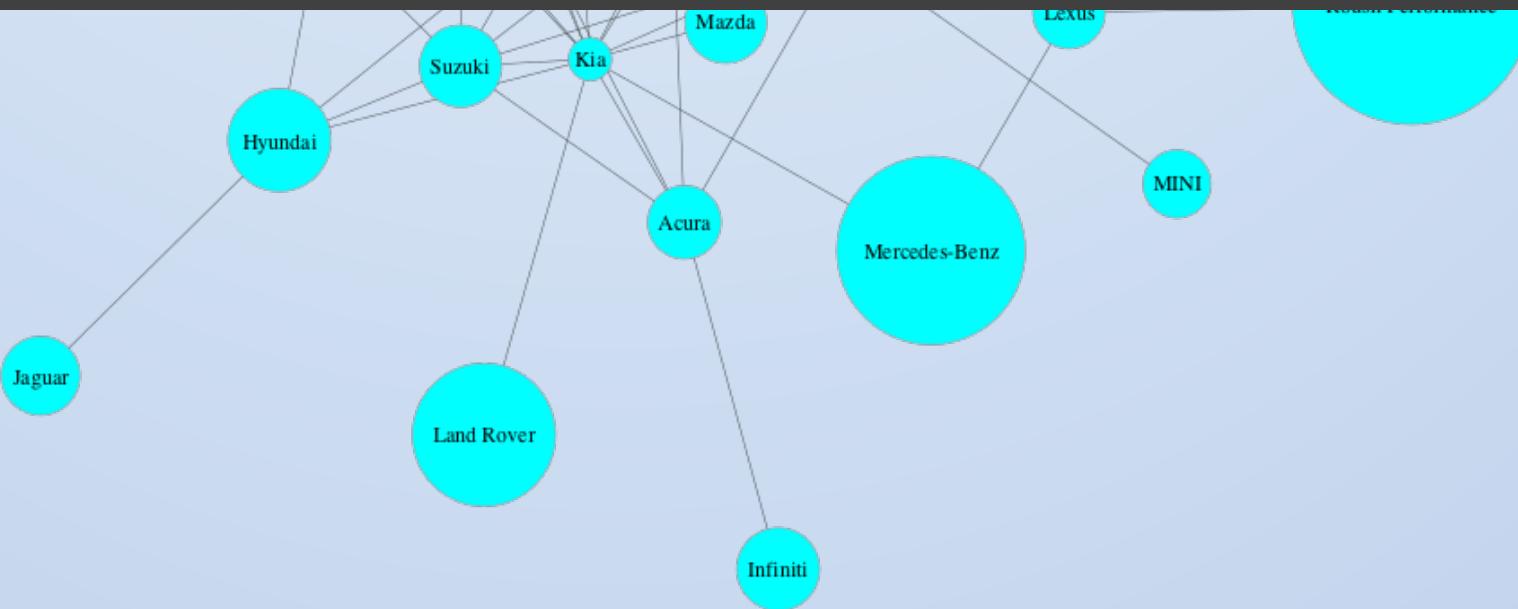
RELATIONAL TO GRAPH TRANSFORMATION

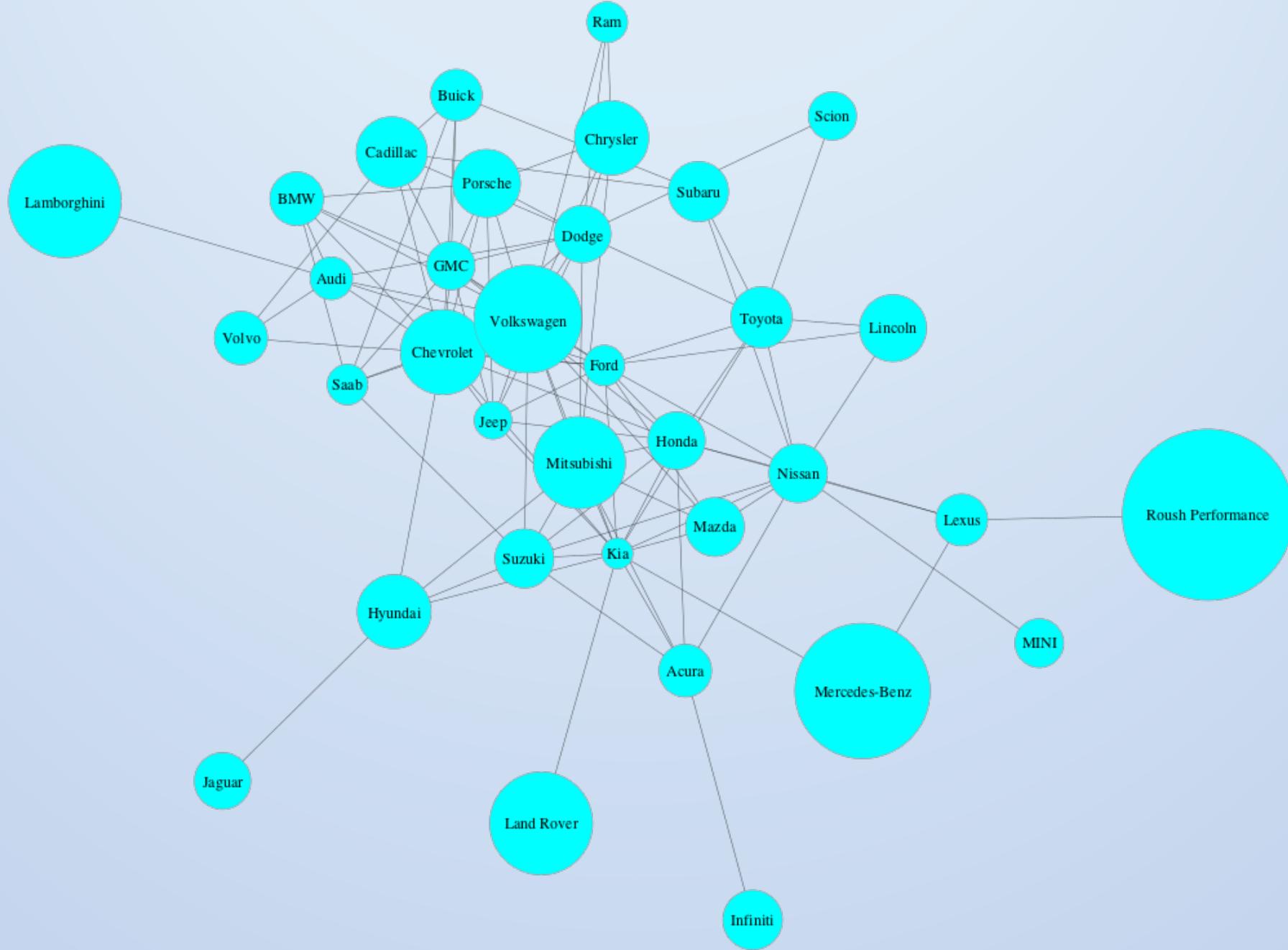
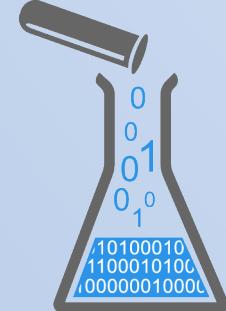
make_x	make_y	edge
Acura	Honda	4
Acura	Infiniti	2
Acura	Lexus	3
Acura	Nissan	1
Acura	Toyota	1
Audi	BMW	7
Audi	Cadillac	1
Audi	Chevrolet	1
Audi	Ford	1
Audi	Hyundai	2

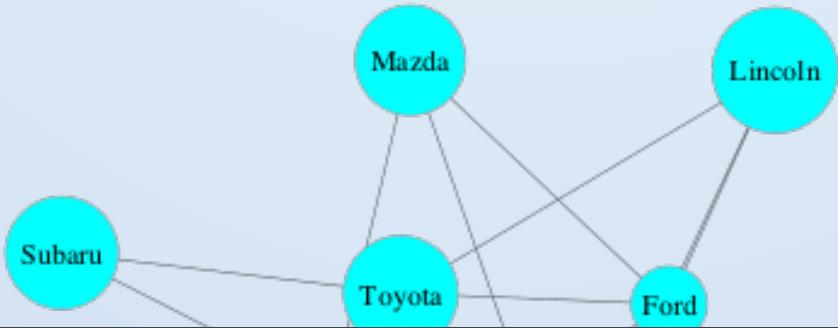




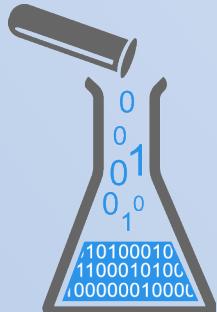
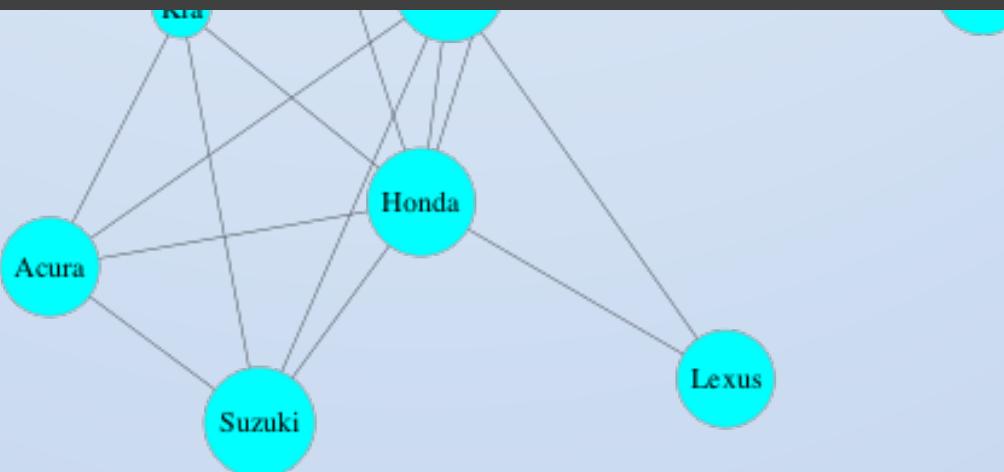
MANUFACTURER NETWORK (2016)

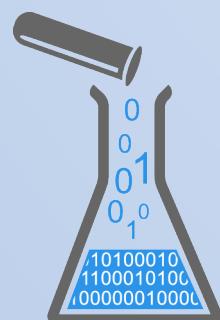
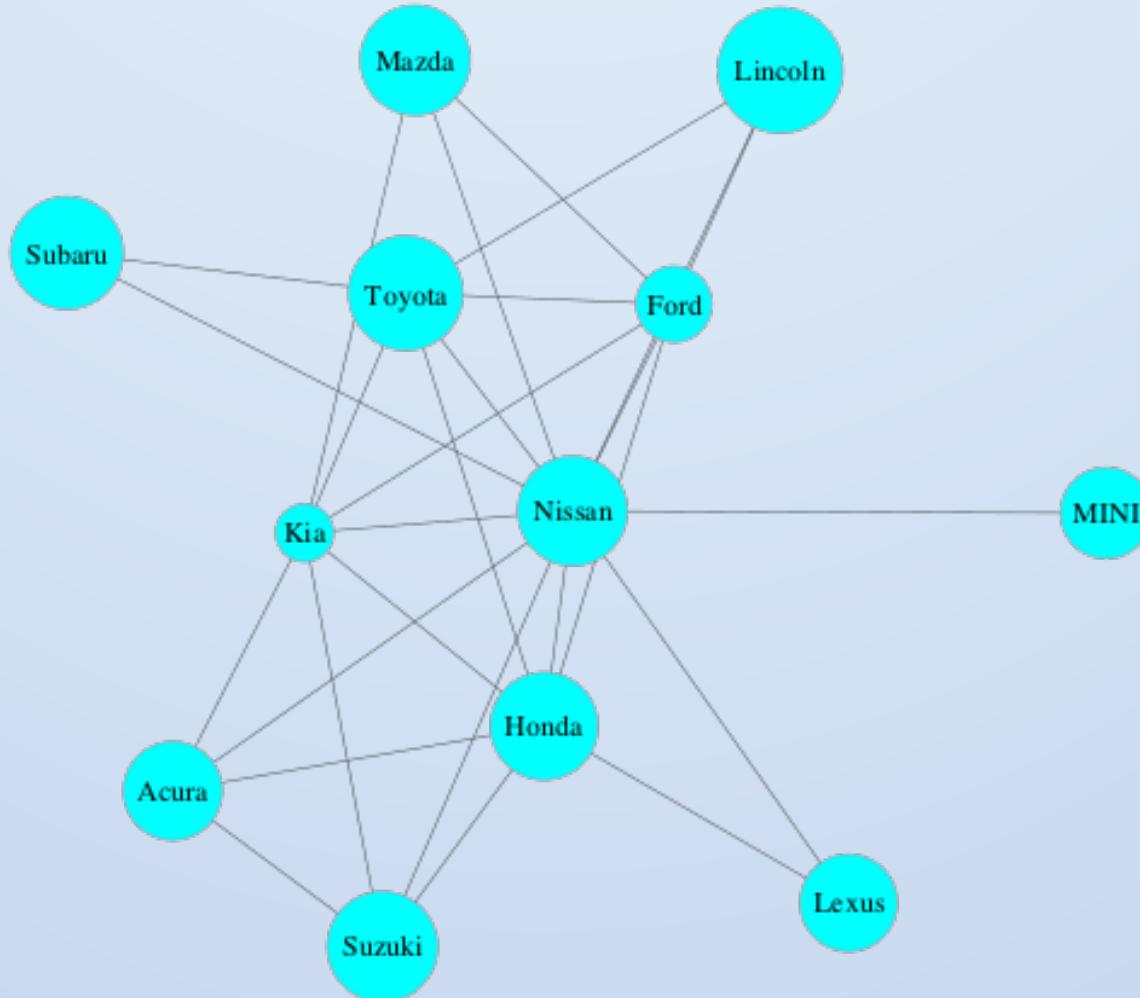


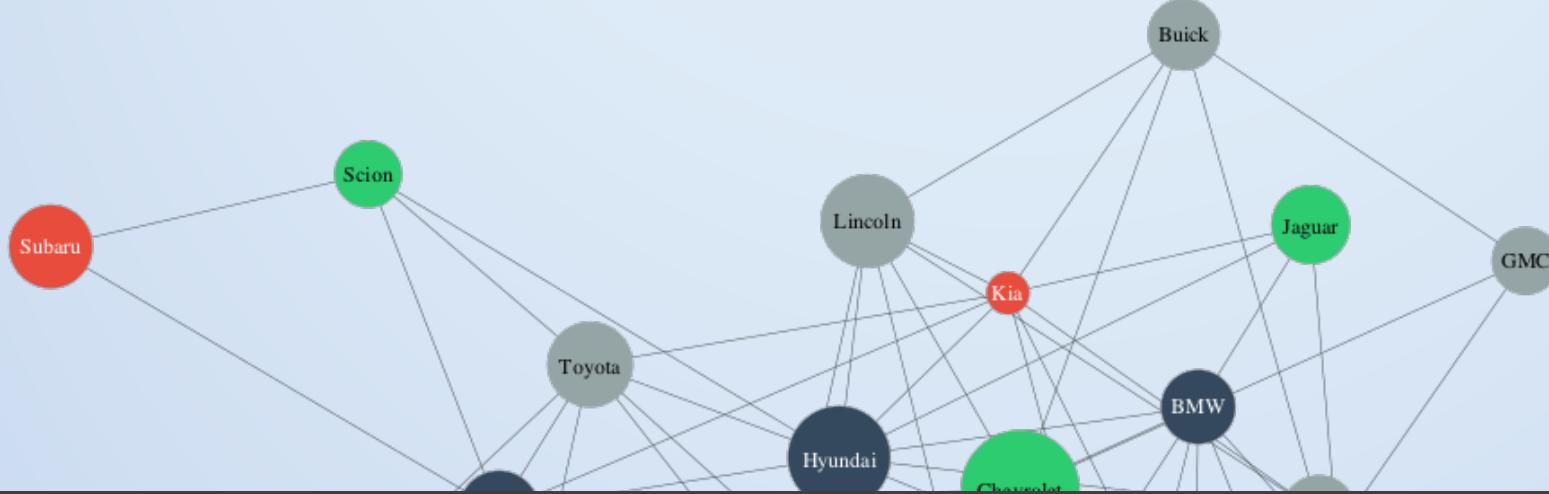




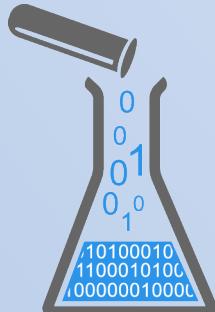
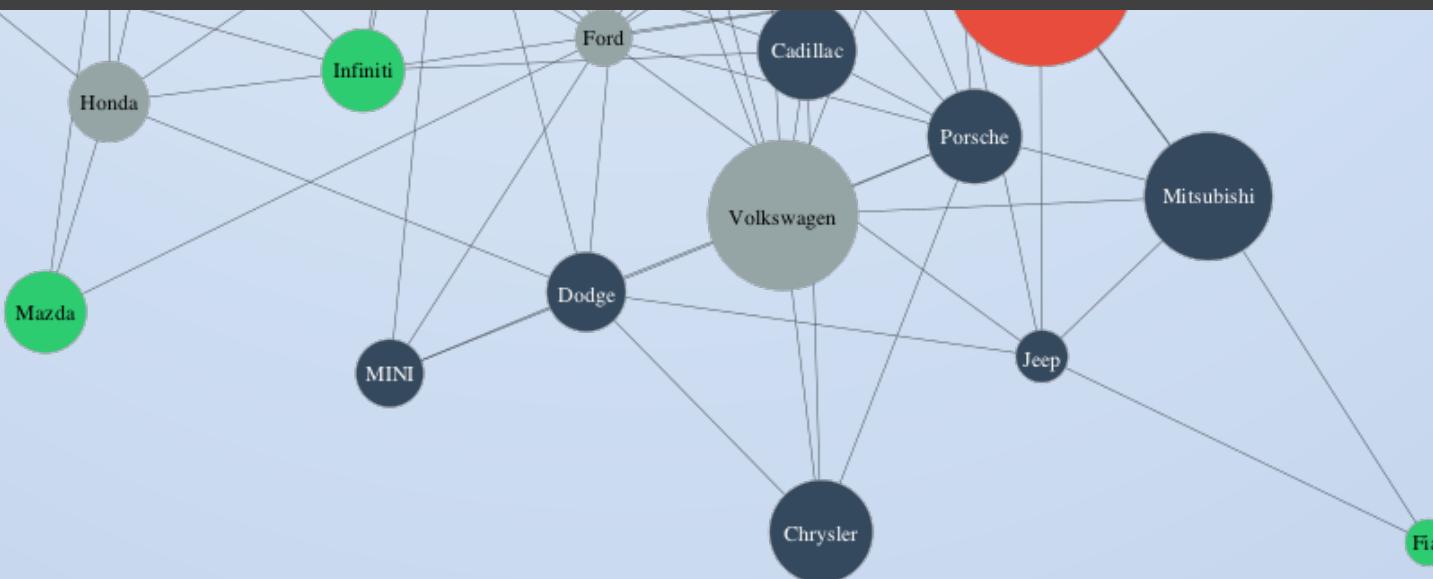
EGO GRAPH FOR NISSAN

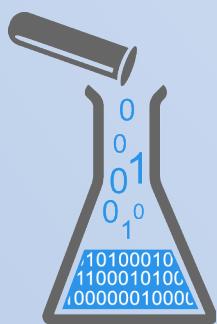


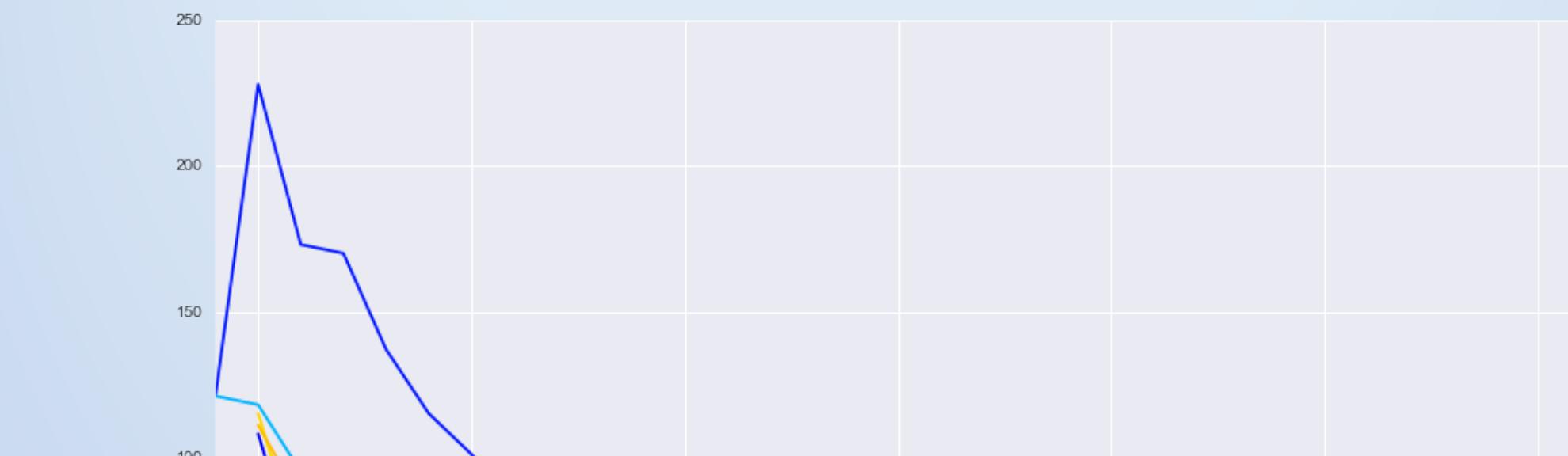




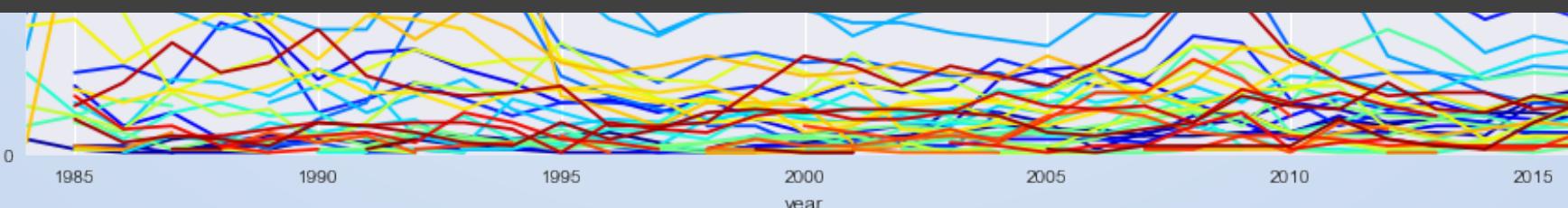
COMMUNITY GRAPH (2016)



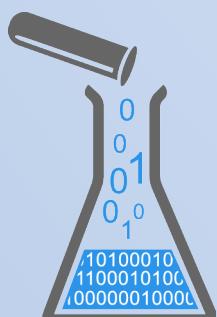


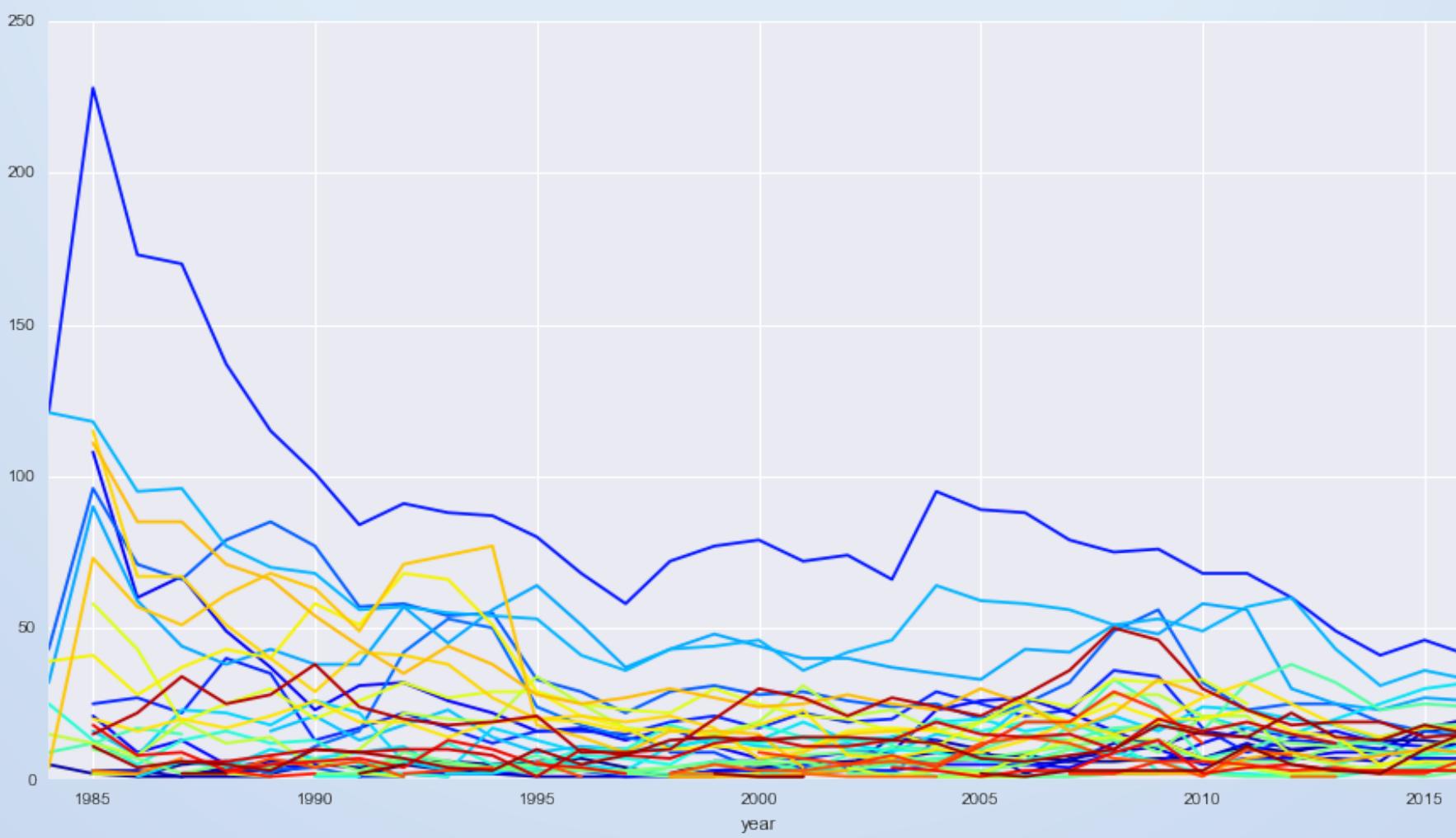


EDGE WEIGHTS OVER TIME

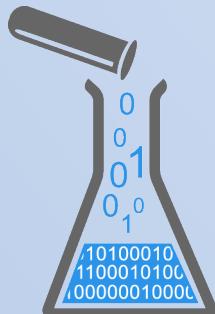


Acura	Daihatsu	J.K. Motors	Mercedes-Benz	S and S Coach Company E.p. Dutton
Alfa Romeo	Dodge	Jaguar	Mercury	Saab
American Motors Corporation	Eagle	Jeep	Merkur	Saleen Performance
Audi	Federal Coach	Kia	Mitsubishi	Saturn
BMW	Ferrari	Lamborghini	Nissan	Scion
Bentley	Fiat	Land Rover	Oldsmobile	Sterling
Bill Dovell Motor Car Company	Ford	Lexus	Pontiac	Subaru
Buick	GMC	Lincoln	Porsche	Superior Coaches Div E.p. Dutton
CCC Engineering	Geo	Lotus	Quantum Technologies	Suzuki
Cadillac	Honda	MINI	Ram	Tecstar, LP
Chevrolet	Hyundai	Maserati	Renault	Toyota
Chrysler	Import Trade Services	Maybach	Rolls-Royce	Volkswagen
Dabryan Coach Builders Inc	Infiniti	Mazda	Roush Performance	Volvo
Daewoo	Isuzu	Mcevoy Motors	Wallace Environmental	





Acura	Daihatsu	J.K. Motors	Mercedes-Benz	S and S Coach Company E.p. Dutton
Alfa Romeo	Dodge	Jaguar	Mercury	Saab
American Motors Corporation	Eagle	Jeep	Merkur	Saleen Performance
Audi	Federal Coach	Kia	Mitsubishi	Saturn
BMW	Ferrari	Lamborghini	Nissan	Scion
Bentley	Fiat	Land Rover	Oldsmobile	Sterling
Bill Dovell Motor Car Company	Ford	Lexus	Pontiac	Subaru
Buick	GMC	Lincoln	Porsche	Superior Coaches Div E.p. Dutton
CCC Engineering	Geo	Lotus	Quantum Technologies	Suzuki
Cadillac	Honda	MINI	Ram	Tecstar, LP
Chevrolet	Hyundai	Maserati	Renault	Toyota
Chrysler	Import Trade Services	Maybach	Rolls-Royce	Volkswagen
Dabryan Coach Builders Inc	Infiniti	Mazda	Roush Performance	Volvo
Daewoo	Isuzu	Mcevoy Motors	Wallace Environmental	

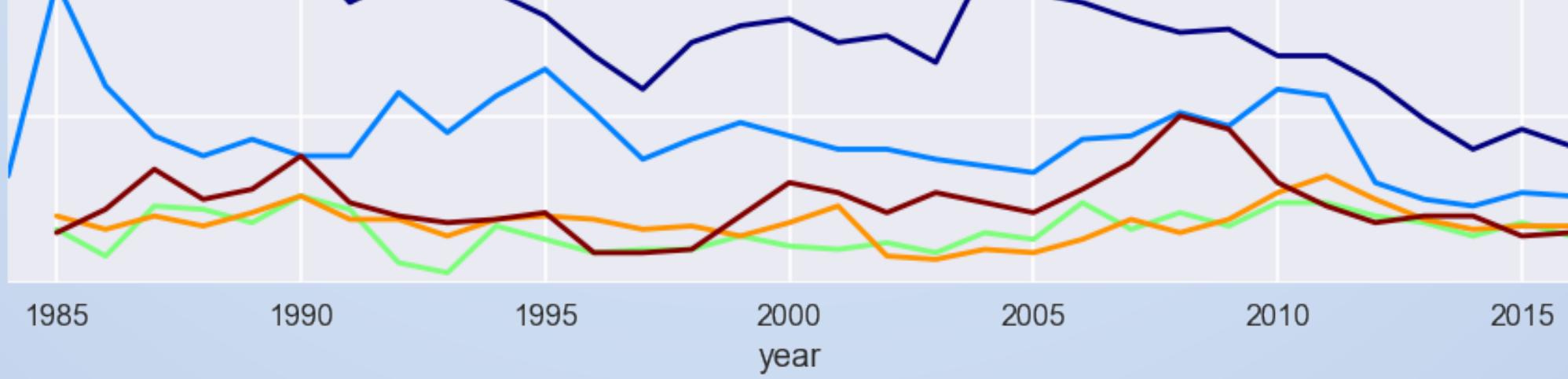


250

200

150

FILTER FOR SPECIFIC MAKES



1985

1990

1995

2000

2005

2010

2015

year

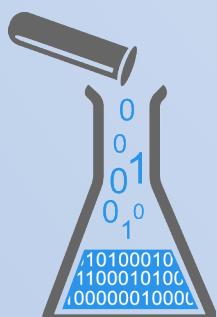
Chevrolet

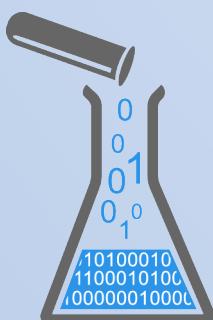
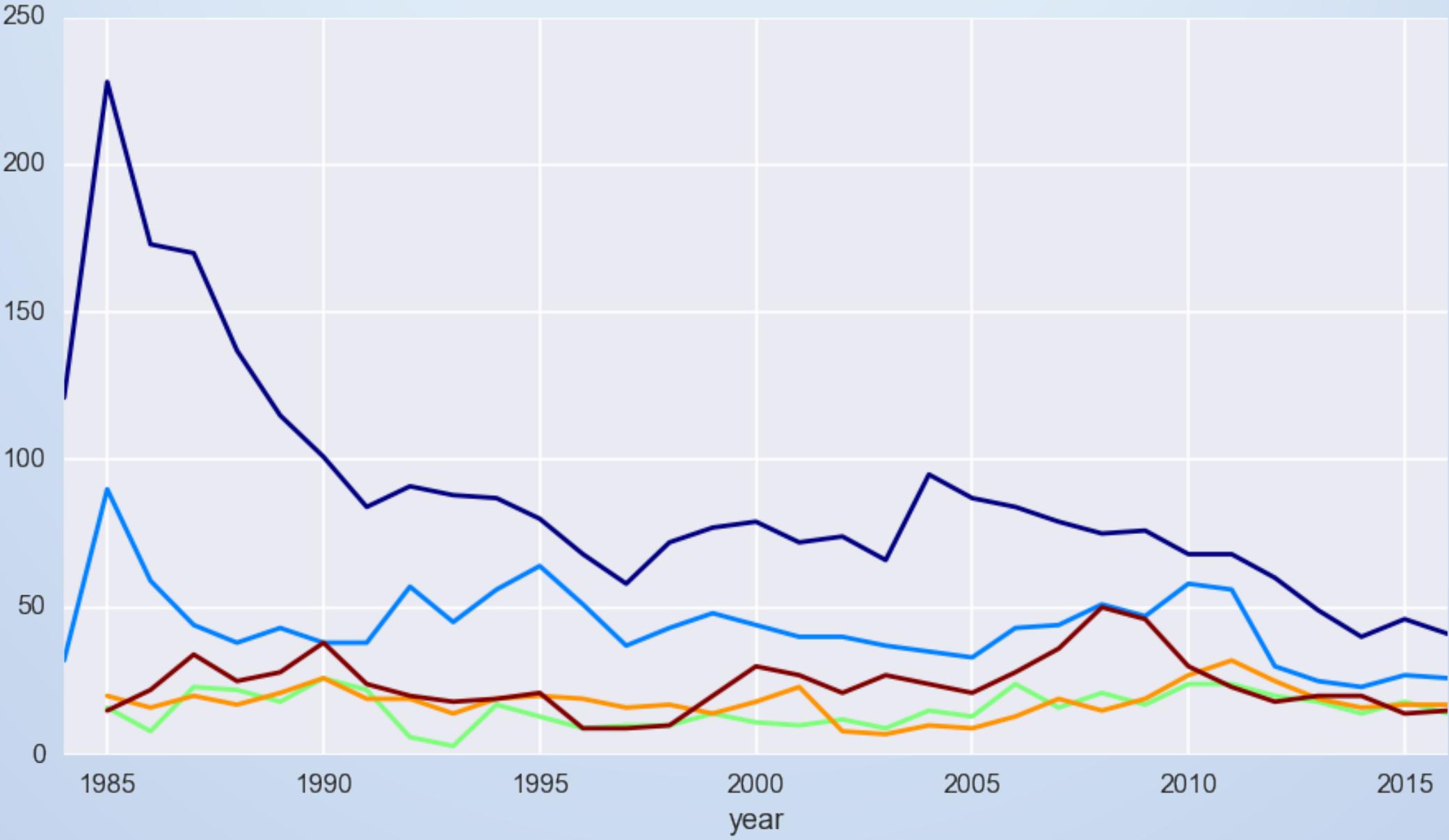
Ford

Honda

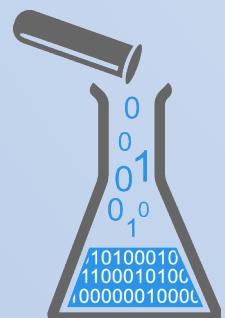
Nissan

Toyota





RECAP



EXPLORATION FRAMEWORK

Prep Phase

Identify

Types of Information

Entities in Data Set

Review

Transformation Methods

Visualization Methods

Create

Category Aggregations

Continuous Bins

Cluster Categories

Explore Phase

Filter + Aggregate

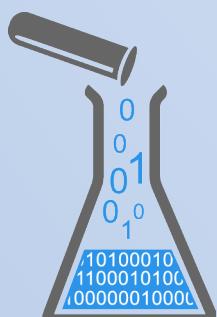
Field Relationships

Entity Relationships

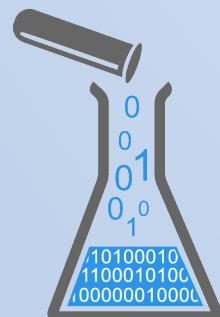
Visualization

Over Time

Insights

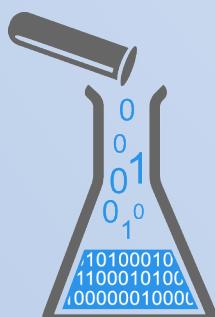


SO MANY INSIGHTS!

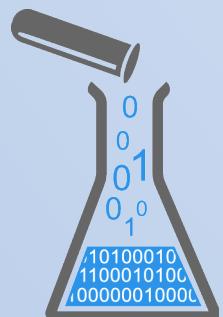


NO REALLY... SO MANY!

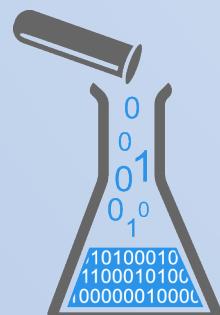
Significantly more small cars than other types in 2016.	Sport Utility Vehicles are currently next most popular.	Midsize Cars currently third most popular.
Other vehicle types currently not as popular.	Sport Utility Vehicles didn't exist in 1985.	Small Cars were even more popular in 1985.
Pickup Trucks were also more popular in 1985.	Special Purpose Vehicles were more popular as well.	Most vehicles today have very small or moderate sized engines.
Most vehicles today are very fuel efficient.	Few vehicles today have large engines.	Even fewer have low fuel efficiency.
BMW currently makes the most vehicle models, followed by Chevy and Ford.	Pagani and Alfa Romeo make the least number of vehicle models.	Smaller cars with smaller engines are most fuel efficient.
Currently no small engine vehicles with low efficiency.	Even Vans and Station Wagons are relatively fuel efficient these days.	Each vehicle category has varying engine sizes.
BMW is doubling down on small cars. So is Porsche.	Ford, Chevrolet, & Nissan are going for breadth.	Jeep and Land Rover are focused solely on SUVs.
Ram is focused solely on Pickup Trucks	A few companies are focused only on small cars.	Toyota used to make a lot of small cars, but now makes less in favor of SUVs and Pickup Trucks.
Overall surge in small efficient engine vehicles over last 10 years.	Mostly at expense of moderate/large inefficient engines.	Even Large Cars are relatively fuel efficient these days.
Linear relationships between engine size and fuel cost and emissions.	Exponential relationships between efficiency and engine size and fuel cost.	Clustering into 4 groups results in relatively clear boundaries in the data.
Manufacturers with both depth and breadth of vehicle attributes have most connections.	Manufacturers that specialize are positioned toward edge of network.	Four distinct communities detected and connections over time have converged.



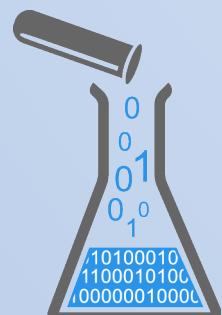
WRAP-UP



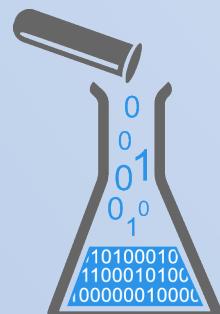
DON'T TAKE DATA
AT FACE VALUE



CAN THIS BE
AUTOMATED?

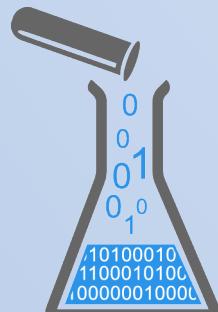


HUMAN+MACHINE COLLABORATION



CULTIWAR

Smart data set management.
Semi-automated, interactive visual exploration.



THANK YOU!

✉ tojeda@districtdatalabs.com

in [linkedin.com/in/tonyojeda](https://www.linkedin.com/in/tonyojeda)

🐦 [@tonyojeda3](https://twitter.com/tonyojeda3)

试管 <http://districtdatalabs.com>

