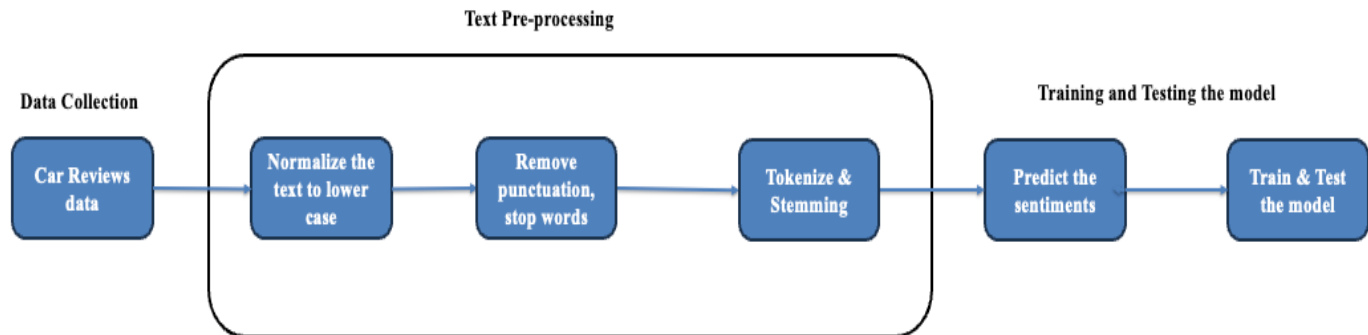


SENTIMENT ANALYSIS PROJECT:

Project Goal: To generate insights from customer feedback, enabling the rental agency to streamline their car inventory based on the sentiments extracted from each review.

Project Flow:



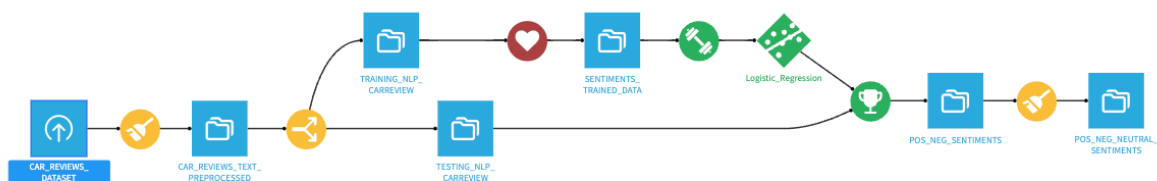
There are 3 main steps:

1. Data Collection
2. Text Pre-processing
3. Training and Testing the model

Data Source: The dataset has been sourced from Kaggle, a popular platform for data science and machine learning projects.

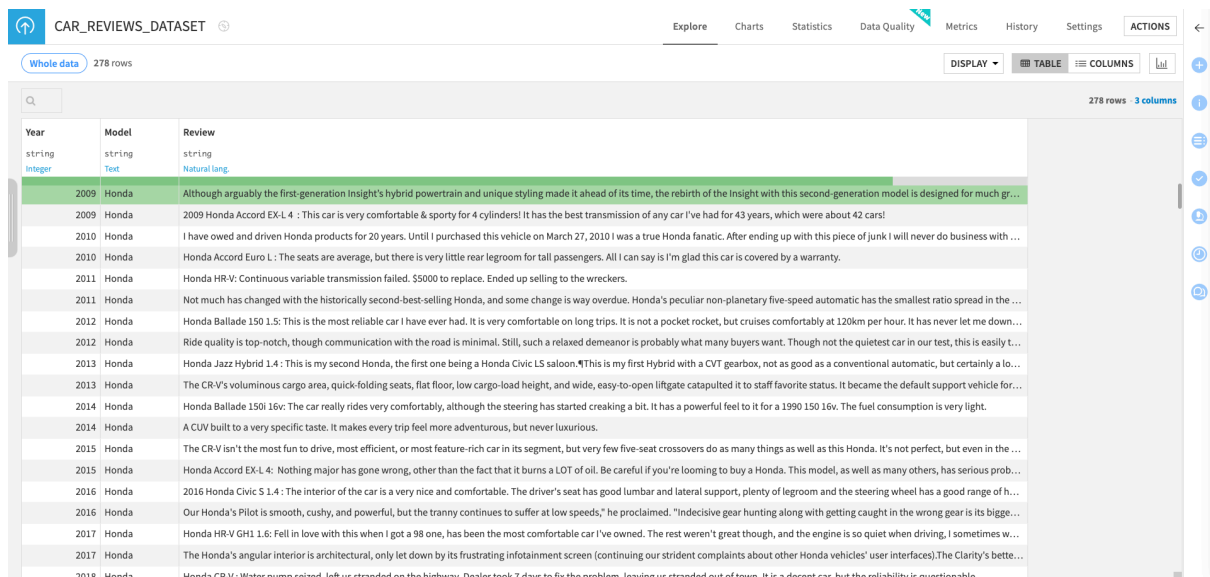
Features:

- **Year:** Year the car was made, providing temporal context to the sentiment analysis.
- **Model:** The model of the car being reviewed, allowing for analysis at a granular level.
- **Reviews:** This feature contains the textual customer reviews, which will be the primary data used to predict sentiments.



Sentiment Analysis Flow in Dataiku

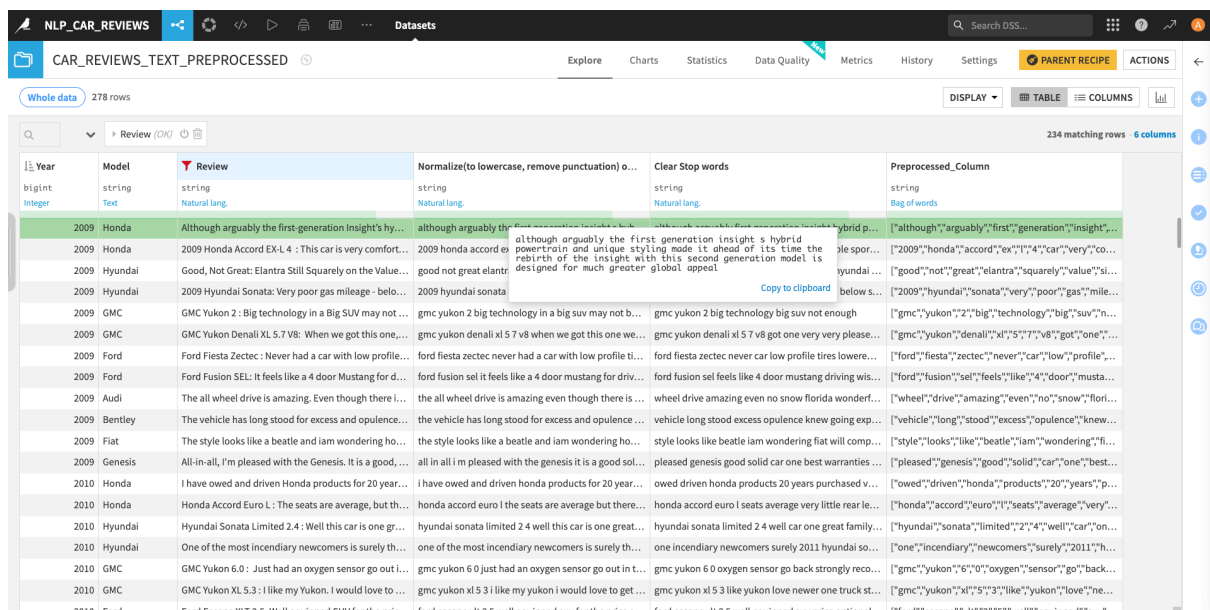
Step1 : Load the Dataset in Dataiku



Year	Model	Review
string Integer	string Text	string Natural lang.
2009	Honda	Although arguably the first-generation Insight's hybrid powertrain and unique styling made it ahead of its time, the rebirth of the Insight with this second-generation model is designed for much gr...
2009	Honda	2009 Honda Accord EX-L 4 : This car is very comfortable & sporty for 4 cylinders! It has the best transmission of any car I've had for 43 years, which were about 42 cars!
2010	Honda	I have owed and driven Honda products for 20 years. Until I purchased this vehicle on March 27, 2010 I was a true Honda fanatic. After ending up with this piece of junk I will never do business with ...
2010	Honda	Honda Accord Euro L: The seats are average, but there is very little rear legroom for tall passengers. All I can say is I'm glad this car is covered by a warranty.
2011	Honda	Honda HR-V: Continuous variable transmission failed. \$5000 to replace. Ended up selling to the wreckers.
2011	Honda	Not much has changed with the historically second-best-selling Honda, and some change is way overdue. Honda's peculiar non-planetary five-speed automatic has the smallest ratio spread in the ...
2012	Honda	Honda Ballade 150 1.5: This is the most reliable car I have ever had. It is very comfortable on long trips. It is not a pocket rocket, but cruises comfortably at 120km per hour. It has never let me down...
2012	Honda	Ride quality is top-notch, though communication with the road is minimal. Still, such a relaxed demeanor is probably what many buyers want. Though not the quietest car in our test, this is easily t...
2013	Honda	Honda Jazz Hybrid 1.4 : This is my second Honda, the first one being a Honda Civic LS saloon.¶This is my first Hybrid with a CVT gearbox, not as good as a conventional automatic, but certainly a lo...
2013	Honda	The CR-V's voluminous cargo area, quick folding seats, flat floor, low cargo-load height, and wide, easy-to-open liftgate catapulted it to staff favorite status. It became the default support vehicle for...
2014	Honda	Honda Ballade 150i 1.6v: The car really rides very comfortably, although the steering has started creaking a bit. It has a powerful feel to it for a 1990 150 1.6v. The fuel consumption is very light.
2014	Honda	A CUV built to a very specific taste. It makes every trip feel more adventurous, but never luxurious.
2015	Honda	The CR-V isn't the most fun to drive, most efficient, or most feature-rich car in its segment, but very few five-seat crossovers do as many things as well as this Honda. It's not perfect, but even in the ...
2015	Honda	Honda Accord EX-L 4: Nothing major has gone wrong, other than the fact that it burns a LOT of oil. Be careful if you're looming to buy a Honda. This model, as well as many others, has serious prob...
2016	Honda	2016 Honda Civic S 1.4 : The interior of the car is a very nice and comfortable. The driver's seat has good lumbar and lateral support, plenty of legroom and the steering wheel has a good range of h...
2016	Honda	Our Honda's Pilot is smooth, cushy, and powerful, but the tranny continues to suffer at low speeds," he proclaimed. "Indecisive gear hunting along with getting caught in the wrong gear is its bigge...
2017	Honda	Honda HR-V GH1 1.6: Fell in love with this when I got a 98 one, has been the most comfortable car I've owned. The rest weren't great though, and the engine is so quiet when driving. I sometimes w...
2017	Honda	The Honda's angular interior is architectural, only let down by its frustrating infotainment screen (continuing our strident complaints about other Honda vehicles' user interfaces).The Clarity's bette...
2018	Honda	Honda CR-V: Water pump seized. left us stranded on the highway. Dealer took 7 days to fix the problem. leaving us stranded out of town. It is a decent car. but the reliability is questionable.

Step2 : Perform Text Pre-processing on the Review Column

1. Convert to Lower Case and remove punctuation



Year	Model	Review	Normalized(to lowercase, remove punctuation) o...	Clear Stop words	Preprocessed_Column
bigint Integer	string Text	string Natural lang.	string Natural lang.	string Natural lang.	string Bag of words
2009	Honda	Although arguably the first-generation Insight's hy...	although arguably th...	although arguably th...	["although","arguably","first","generation","insight"...]
2009	Honda	2009 Honda Accord EX-L 4 : This car is very comfort...	2009 honda accord ex...	2009 honda accord ex...	["2009","honda","accord","ex","l","4","car","very","co...
2009	Hyundai	Good, Not Great: Elantra Still Squarely on the Value...	good not great elantr...	good not great elantr...	["good","not","great","elantra","squarely","value","si...
2009	Hyundai	2009 Hyundai Sonata: Very poor gas mileage - belo...	2009 hyundai sonata...	2009 hyundai sonata...	["2009","hyundai","sonata","very","poor","gas","mile...
2009	GMC	GMC Yukon 2 : Big technology in a Big SUV may not...	gmc yukon 2 big technology in a big suv may not b...	gmc yukon 2 big technology big suv not enough	["gmc","yukon","2","big","technology","big","suv","n...
2009	GMC	GMC Yukon Denali XL 5.7 V8: When we got this one...	gmc yukon denali xl 5 7 v8 when we got this one we...	gmc yukon denali xl 5 7 v8 got one very very please...	["gmc","yukon","denali","xl","5","7","v8","got","one","...
2009	Ford	Ford Fiesta Zetec: Never had a car with low profile...	ford fiesta zetec never had a car with low profile ti...	ford fiesta zetec never car low profile tires lowere...	["ford","fiesta","zetec","never","car","low","profile",...
2009	Ford	Ford Fusion SEL: It feels like a 4 door Mustang for d...	ford fusion sel it feels like a 4 door mustang for driv...	ford fusion sel feels like 4 door mustang driving wis...	["ford","fusion","sel","feels","like","4","door","musta...
2009	Audi	The all wheel drive is amazing. Even though there i...	the all wheel drive is amazing even though there is ...	wheel drive amazing even no snow florida wonderf...	["wheel","drive","amazing","even","no","snow","flori...
2009	Bentley	The vehicle has long stood for excess and opulence...	the vehicle has long stood for excess and opulence...	vehicle long stood excess opulence knew going exp...	["vehicle","long","stood","excess","opulence","knew...
2009	Fiat	The style looks like a beetle and iam wondering ho...	the style looks like a beetle and iam wondering ho...	style looks like beetle iam wondering fiat will comp...	["style","looks","like","beetle","iam","wondering","fi...
2009	Genesis	All-in-all, I'm pleased with the Genesis. It is a good...	all in all i m pleased with the genesis it is a good sol...	pleased genesis good solid car one best warranties ...	["pleased","genesis","good","solid","car","one","best...
2010	Honda	I have owed and driven Honda products for 20 year...	i have owed and driven honda products for 20 year...	owed driven honda products 20 years purchased v...	["owed","driven","honda","products","20","years","p...
2010	Honda	Honda Accord Euro L: The seats are average, but th...	honda accord euro l the seats are average but there...	honda accord euro l seats average very little rear le...	["honda","accord","euro","l","seats","average","very"...]
2010	Hyundai	Hyundai Sonata Limited 2.4 : Well this car is one gr...	hyundai sonata limited 2 4 well this car is one great...	hyundai sonata limited 2 4 well car one great family...	["hyundai","sonata","limited","2","4","well","car","on...
2010	Hyundai	One of the most incendiary newcomers is surely th...	one of the most incendiary newcomers is surely th...	one incendiary newcomers surely 2011 hyundai so...	["one","incendiary","newcomers","surely","2011","h...
2010	GMC	GMC Yukon 6.0 : Just had an oxygen sensor go out i...	gmc yukon 6 0 just had an oxygen sensor go out in t...	gmc yukon 6 0 oxygen sensor go back strongly reco...	["gmc","yukon","6","0","oxygen","sensor","go","back...
2010	GMC	GMC Yukon XL 5.3 : I like my Yukon. I would love to ...	gmc yukon xl 5 3 i like my yukon i would love to get ...	gmc yukon xl 5 3 like yukon love newer one truck st...	["gmc","yukon","xl","5","3","like","yukon","love","ne...
2010	Ford	Ford Escape XLT 2.5: Well equinned SUV for the pric...	ford escape xlt 2 5 well equinned suv for the price a...	ford escape xlt 2 5 well equinned suv price optional...	["ford","escape","xlt","2","5","well","equinned","suv"...]

I/P: "The car was GREAT! It exceeded my expectations, and I had an awesome experience!!!!"

- **Convert to Lower Case:**
 - Before: "The car was GREAT! It exceeded my expectations, and I had an awesome experience!!!!"
 - After: "the car was great! it exceeded my expectations, and i had an awesome experience!!!!"

- **Remove Punctuation:**

- Before: "the car was great! it exceeded my expectations, and i had an awesome experience!!!"
- After: "the car was great it exceeded my expectations and i had an awesome experience"

2. Clear stop words : To remove common words that do not carry significant meaning and are not useful for sentiment analysis. These words, known as stop words, include terms like "the," "is," "in," "and," etc. In python, we use NLP libraries such as NLTK

Year	Model	Review	Normalized(to lowercase, remove punctuation) o...	Clear Stop words	Preprocessed_Column
2009	Honda	Although arguably the first-generation Insight's hy...	although arguably the first generation insight's hyb...	although arguab...	["although","arguab...","first","generation","insight"...]
2009	Honda	2009 Honda Accord EX-L 4 : This car is very comfort...	2009 honda accord ex l 4 this car is very comfortabl...	2009 honda	["2009","honda","accord","ex","l","4","this","car","is","very","comfortabl..."]
2009	Hyundai	Good, Not Great: Elantra Still Squarely on the Value...	good not great elantra still squarely on the value sl...	good not gr	["good","not","great","elantra","still","squarely","on","the","value","sl..."]
2009	Hyundai	2009 Hyundai Sonata: Very poor gas mileage - belo...	2009 hyundai sonata very poor gas mileage below s...	2009 hyund	["2009","hyundai","sonata","very","poor","gas","mileage","below","s..."]
2009	GMC	GMC Yukon 2 : Big technology in a Big SUV may not...	gmc yukon 2 big technology in a big suv may not b...	gmc yukon 2 big technology big suv not enough	["gmc","yukon","2","big","technology","in","a","big","suv","may","not","b..."]
2009	GMC	GMC Yukon Denali XL 5.7 V8: When we got this one...	gmc yukon denali xl 5 7 v8 when we got this one we...	gmc yukon denali xl 5 7 v8 got one very very please...	["gmc","yukon","denali","xl","5","7","v8","got","one","very","very","please..."]
2009	Ford	Ford Fiesta Zetec: Never had a car with low profile...	ford fiesta zetec never had a car with low profile ti...	ford fiesta zetec never car low profile tires lowere...	["ford","fiesta","zetec","never","car","low","profile","tires","lowere..."]
2009	Ford	Ford Fusion SEL: It feels like a 4 door Mustang for d...	ford fusion sel it feels like a 4 door mustang for driv...	ford fusion sel feels like 4 door mustang driving wis...	["ford","fusion","sel","it","feels","like","a","4","door","mustang","for","driv..."]
2009	Audi	The all wheel drive is amazing. Even though there i...	the all wheel drive is amazing even though there is ...	wheel drive amazing even no snow florida wonderf...	["wheel","drive","is","amazing","even","though","there","is","a","good","sol..."]
2009	Bentley	The vehicle has long stood for excess and opulence...	the vehicle has long stood for excess and opulence ...	vehicle long stood excess opulence knew going exp...	["vehicle","long","stood","for","excess","and","opulence","knew","going","exp..."]
2009	Fiat	The style looks like a beetle and iam wondering ho...	the style looks like a beetle and iam wondering ho...	style looks like beetle iam wondering fiat will comp...	["style","looks","like","a","beetle","and","iam","wondering","ho..."]
2009	Genesis	All-in-all, I'm pleased with the Genesis. It is a good...	all in all i m pleased with the genesis it is a good sol...	pleased genesis good solid car one best warranties ...	["pleased","genesis","good","solid","car","one","best","warranties","..."]
2010	Honda	I have owed and driven Honda products for 20 year...	i have owed and driven honda products for 20 year...	owed driven honda products 20 years purchased v...	["owed","driven","honda","products","for","20","years","purchased","v..."]
2010	Honda	Honda Accord Euro L : The seats are average, but th...	honda accord euro l the seats are average but there...	honda accord euro l seats average very little rear le...	["honda","accord","euro","l","the","seats","are","average","but","th..."]
2010	Hyundai	Hyundai Sonata Limited 2.4 : Well this car is one gr...	hyundai sonata limited 2 4 well this car is one great...	hyundai sonata limited 2 4 well car one great family...	["hyundai","sonata","limited","2","4","well","this","car","is","one","gr..."]
2010	Hyundai	One of the most incendiary newcomers is surely th...	one of the most incendiary newcomers is surely th...	one incendiary newcomers surely 2011 hyundai so...	["one","of","the","most","incendiary","newcomers","is","surely","th..."]
2010	GMC	GMC Yukon 6.0 : Just had an oxygen sensor go out i...	gmc yukon 6 0 just had an oxygen sensor go out in t...	gmc yukon 6 0 oxygen sensor go back strongly reco...	["gmc","yukon","6","0","just","had","an","oxygen","sensor","go","out","i..."]
2010	GMC	GMC Yukon XL 5.3 : I like my Yukon. I would love to ...	gmc yukon xl 5 3 i like my yukon i would love to get ...	gmc yukon xl 5 3 like yukon love newer one truck st...	["gmc","yukon","xl","5","3","i","like","my","yukon","i","would","love","to","get..."]
2010	Ford	Ford Escape XLT 2.5: Well equipped SUV for the pric...	ford escape xlt 2 5 well equipped suv for the price a...	ford escape xlt 2 5 well equipped suv price optional...	["ford","escape","xlt","2","5","well","equipped","suv","for","the","price..."]

Before: "the car was great it exceeded my expectations and i had an awesome experience"

After: "car great exceeded expectations awesome experience"

3. Stemming and Tokenizing :

Stemming : Reduce words to their base or root form, which helps in consolidating different variations of the same word.

For example, "running," "runs," and "ran" can all be reduced to "run."

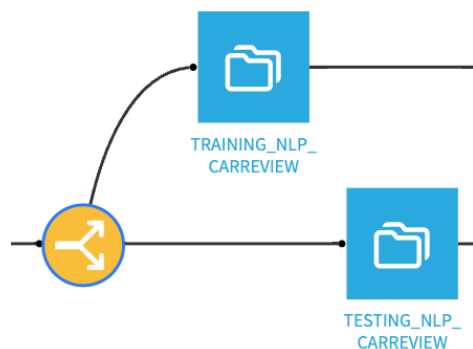
Tokenizing : To split the text into individual words or tokens, which is a crucial step for further analysis.

Example: "The car was great" becomes ["the", "car", "was", "great"]

NLP_CAR_REVIEWS					
CAR_REVIEWS_TEXT_PREPROCESSED					
278 rows					
Year	Model	Review	Normalized(to lowercase, remove punctuation) o...	Clear Stop words	Preprocessed_Column
bigint	string	string	string	string	string
integer	Text	Natural lang.	Natural lang.	Natural lang.	Bag of words
2009	Honda	Although arguably the first-generation Insight's hy...	although arguably the first generation insight s hyb...	although areuablv first generation insieght hvrbrid o...	["although","areuablv","first","generation","insight",...
2009	Honda	2009 Honda Accord EX-L 4 : This car is very comfort...	2009 honda accord ex l 4 this car is very comfortabl...	[["accord","ex","l","4","car","very","co...]
2009	Hyundai	Good, Not Great: Elantra Still Squarely on the Value...	good not great elantra still squarely on the value si...	["although",	["t","elantra","squarely","value","si...]
2009	Hyundai	2009 Hyundai Sonata: Very poor gas mileage - belo...	2009 hyundai sonata very poor gas mileage below s...	["arguably",	["sonata","very","poor","gas","mile...]
2009	GMC	GMC Yukon 2 : Big technology in a Big SUV may not ...	gmc yukon 2 big technology in a big suv may not b...	["first",	["big","technology","big","suv","n...]
2009	GMC	GMC Yukon Denali XL 5.7 V8: When we got this one...	gmc yukon denali xl 5 7 v8 when we got this one we...	["hybrid",	["hali","xl","5","7","v8","got","one",...
2009	Ford	Ford Fiesta Zetec: Never had a car with low profile...	ford fiesta zetec never had a car with low profile ti...	["powertrain",	["tec","never","car","low","profile",...
2009	Ford	Ford Fusion SEL: It feels like a 4 door Mustang for d...	ford fusion sel it feels like a 4 door mustang for driv...	["unique",	["","feels","like","4","door","musta...]
2009	Audi	The all wheel drive is amazing. Even though there i...	the all wheel drive is amazing even though there is ...	["styling",	["hazing","even","no","snow","flori...]
2009	Bentley	The vehicle has long stood for excess and opulence...	the vehicle has long stood for excess and opulence ...	["node",	["ood","excess","opulence","knew...]
2009	Fiat	The style looks like a beetle and iam wondering ho...	the style looks like a beetle and iam wondering ho...	["ahead",	["style","looks","like","beetle","iam","wondering","fi...]
2009	Genesis	All-in-all, I'm pleased with the Genesis. It is a good...	all in all i m pleased with the genesis it is a good sol...	["time",	["pleased","genesis","good","solid","car","one","best...]
2010	Honda	I have owed and driven Honda products for 20 year...	i have owed and driven honda products for 20 year...	["rebirth",	["owed","driven","honda","products","20","years","p...]
2010	Honda	Honda Accord Euro L: The seats are average, but th...	honda accord euro l the seats are average but there...	Array with elements (chars)	["honda","accord","euro","l","seats","average","very...]
2010	Hyundai	Hyundai Sonata Limited 2.4 : Well this car is one gr...	hyundai sonata limited 2 4 well this car is one great...	Copy to clipboard View raw value	["hyundai","sonata","limited","2","4","well","car","on...]
2010	Hyundai	One of the most incendiary newcomers is surely th...	one of the most incendiary newcomers is surely th...		["one","incendiary","newcomers","surely","2011","h...]
2010	GMC	GMC Yukon 6.0 : Just had an oxygen sensor go out i...	gmc yukon 6 0 just had an oxygen sensor go out in t...		["gmc","yukon","6","0","oxygen","sensor","go","back...]
2010	GMC	GMC Yukon XL 5.3 : I like my Yukon. I would love to ...	gmc yukon xl 5 3 i like my yukon i would love to get ...		["gmc","yukon","xl","5","3","like","yukon","love","ne...]
2010	Ford	Ford Escape XLT 2.5: Well equinned SUV for the pric...	ford escape xlt 2 5 well equinned suv for the price a...		["ford","escape","xlt","2","5","well","equinned","suv",...

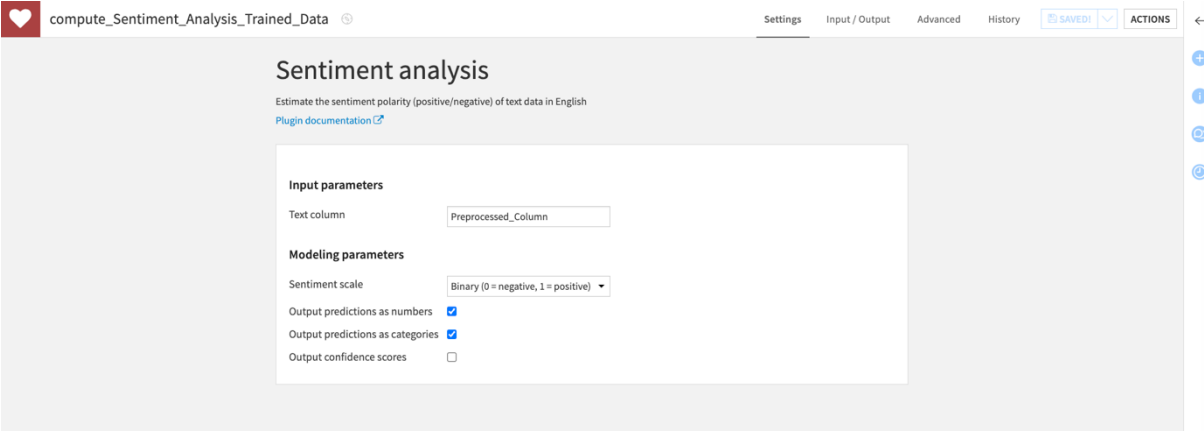
Pre-processed Tokenized Review column

Then, Splitted the data into two parts: 70% for training and 30% for testing.



- **Training Data:** The subset of data used to train machine learning models, enabling them to learn patterns and relationships from the input features and their corresponding labels.
- **Testing Data:** The subset of data held out from the training process, used to evaluate the trained model's performance on unseen examples to assess its predictive accuracy and generalization ability.

I used the Sentiment Analysis Plug-in in Dataiku to predict sentiments on training data. Since the training data doesn't have corresponding labels such as positive or negative, I explored its functionalities and learned how to effectively apply it within the platform.

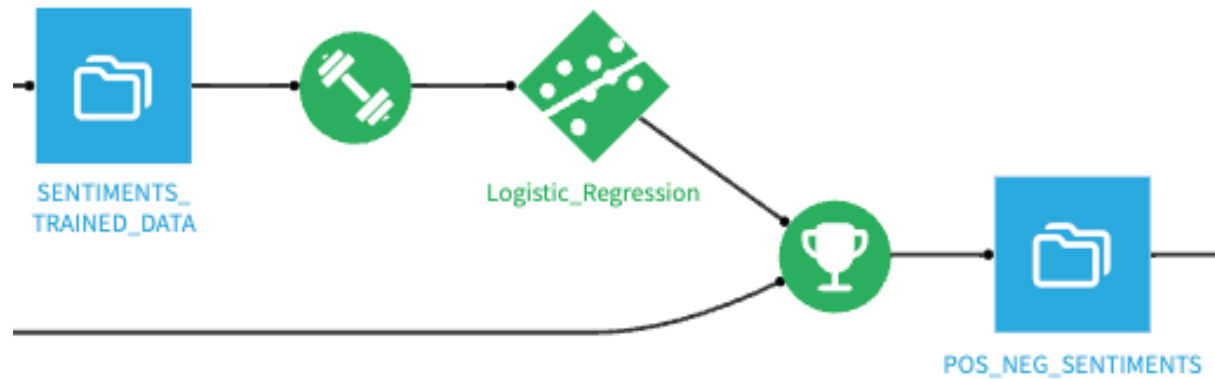


Then, I predicted sentiments on the training data. The image below highlights the sentiment predicted as Positive.

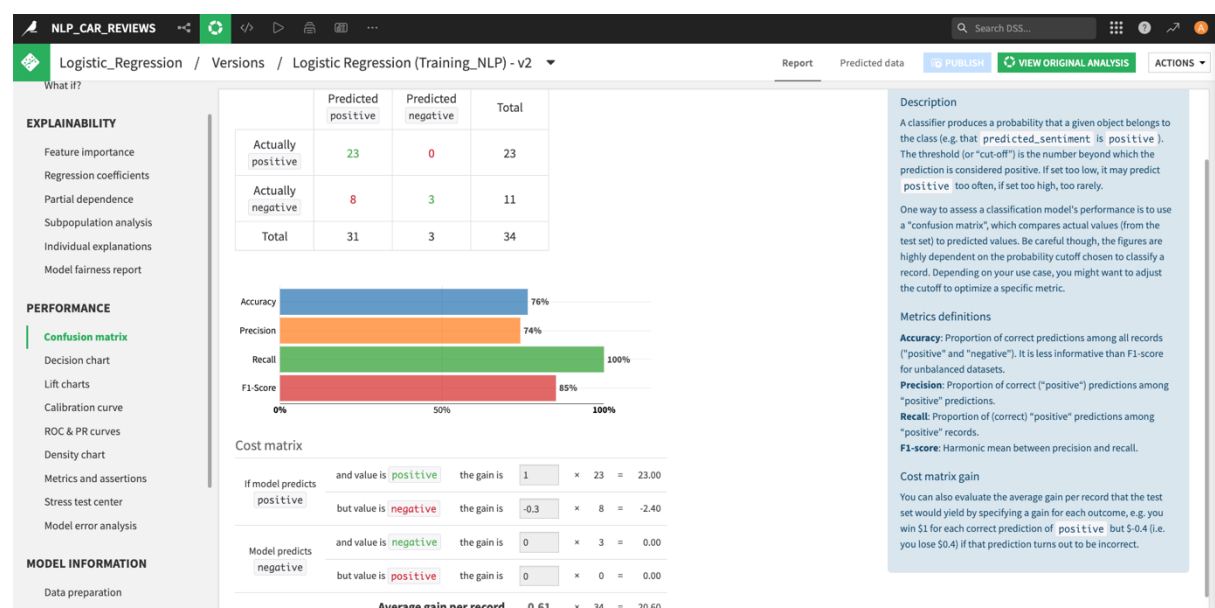
The screenshot shows the 'SENTIMENTS_TRAINED_DATA' dataset in Dataiku. It displays 174 rows. The first row is highlighted in green, indicating a positive sentiment. The columns are: Review, Normalize(to lowercase, remove punctuation) o..., Clear Stop words, Preprocessed_Column, predicted_score, and predicted_sentiment. The 'predicted_sentiment' column shows 'positive' for the highlighted row.

Review	Normalize(to lowercase, remove punctuation) o...	Clear Stop words	Preprocessed_Column	predicted_score	predicted_sentiment
ough arguably the first-generation insight's hy...	although arguably the first generation insight's hyb...	although arguably first generation insight hybrid p...	["although","arguably","first","generation","insight"...	1	positive
9 Honda Accord EX-L 4 : This car is very comfort...	2009 honda accord ex l 4 this car is very comfortabl...	2009 honda accord ex l 4 car very comfortable spor...	["2009","honda","accord","ex","l","4","car","very","comfortable","spor..."	0	negative
jd, Not Great: Elantra Still Squarely on the Value...	good not great elantra still squarely on the value si...	good not great elantra squarely value side hyundai ...	["good","not","great","elantra","squarely","value","side","hyundai"...	0	negative
9 Hyundai Sonata: Very poor gas mileage - belo...	2009 hyundai sonata very poor gas mileage below s...	2009 hyundai sonata very poor gas mileage below s...	["2009","hyundai","sonata","very","poor","gas","mileage","below","s..."	0	negative
C Yukon 2 : Big technology in a Big SUV may not ...	gmc yukon 2 big technology in a big suv may not b...	gmc yukon 2 big technology big suv not enough	["gmc","yukon","2","big","technology","big","suv","not","enough"	0	negative
C Yukon Denali XL 5.7 V8: When we got this one...	gmc yukon denali xl 5 7 v8 when we got this one we...	gmc yukon denali xl 5 7 v8 got one very very please...	["gmc","yukon","denali","xl","5","7","v8","got","one","very","very","please..."	1	positive
d Fiesta Zectec : Never had a car with low profile...	ford fiesta zectec never had a car with low profile ti...	ford fiesta zectec never car low profile tires lowere...	["ford","fiesta","zectec","never","car","low","profile","tires","lowere..."	0	negative
d Fusion SEL: It feels like a 4 door Mustang for driv...	ford fusion sel it feels like a 4 door mustang for driv...	ford fusion sel feels like 4 door mustang driving wis...	["ford","fusion","sel","feels","like","4","door","mustang","driving","wis..."	1	positive
: all wheel drive is amazing. Even though there i...	the all wheel drive is amazing even though there is ...	wheel drive amazing even no snow florida wonderf...	["wheel","drive","amazing","even","no","snow","florida","wonderful..."	1	positive
: vehicle has long stood for excess and opulence...	the vehicle has long stood for excess and opulence ...	vehicle long stood excess opulence knew going exp...	["vehicle","long","stood","excess","opulence","knew","going","expensive..."	1	positive
: style looks like a beetle and iam wondering ho...	the style looks like a beetle and iam wondering ho...	style looks like beetle iam wondering fiat will comp...	["style","looks","like","beetle","iam","wondering","fiat","will","comp..."	1	positive
in-all, I'm pleased with the Genesis. It is a good...	all in all i m pleased with the genesis it is a good sol...	pleased genesis good solid car one best warranties ...	["pleased","genesis","good","solid","car","one","best","warranties","..."	1	positive
ve owed and driven Honda products for 20 year...	i have owed and driven honda products for 20 year...	owed driven honda products 20 years purchased v...	["owed","driven","honda","products","20","years","purchased","very..."	0	negative
nda Accord Euro L : The seats are average, but th...	honda accord euro l the seats are average but there...	honda accord euro l seats average very little rear le...	["honda","accord","euro","l","seats","average","very","little","rear","le..."	0	negative
ndal Sonata Limited 2.4 : Well this car is one gr...	hyundai sonata limited 2 4 well this car is one great...	hyundai sonata limited 2 4 well car one great family...	["hyundai","sonata","limited","2","4","well","car","one","great","family..."	1	positive
: one of the most incendiary newcomers is surely th...	one of the most incendiary newcomers is surely th...	one incendiary newcomers surely 2011 hyundai so...	["one","incendiary","newcomers","surely","2011","hyundai","sonata","limited..."	1	positive
C Yukon 6.0 : Just had an oxygen sensor go out i...	gmc yukon 6 0 just had an oxygen sensor go out in t...	gmc yukon 6 0 oxygen sensor go back strongly reco...	["gmc","yukon","6","0","oxygen","sensor","go","out","in","the","truck","..."	1	positive
C Yukon XL 5.3 : I like my Yukon. I would love to ...	gmc yukon xl 5 3 i like my yukon i would love to get ...	gmc yukon xl 5 3 like yukon love newer one truck st...	["gmc","yukon","xl","5","3","like","yukon","love","newer","one","truck","st..."	1	positive
d Escape XLT 2.5: Well equipped SUV for the price...	ford escape xlt 2 5 well equipped suv for the price a...	ford escape xlt 2 5 well equipped suv price rational...	["ford","escape","xlt","2","5","well","equipped","suv","for","the","price","a..."	0	negative

Deployed the Logistic regression model and trained on Sentiments_Trained_Data, then used this model to test on Unseen Testing Data

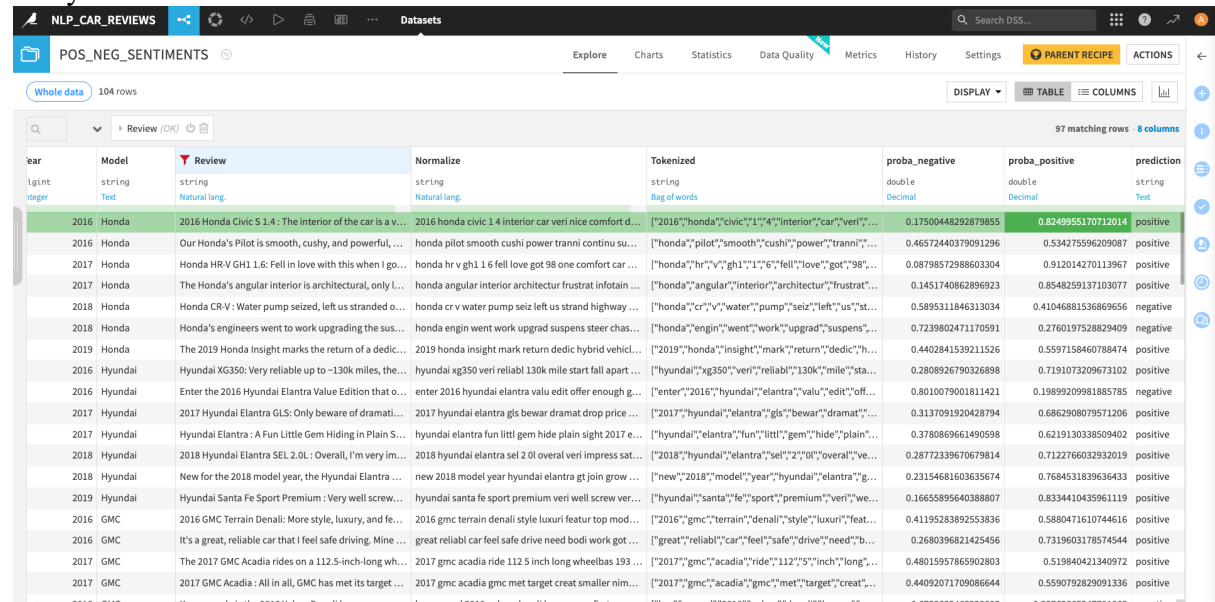


In the initial iteration, the model achieved an accuracy of 76%. I further assessed its performance using additional metrics, such as the confusion matrix, detailed below:



Positive and Negative Sentiments:

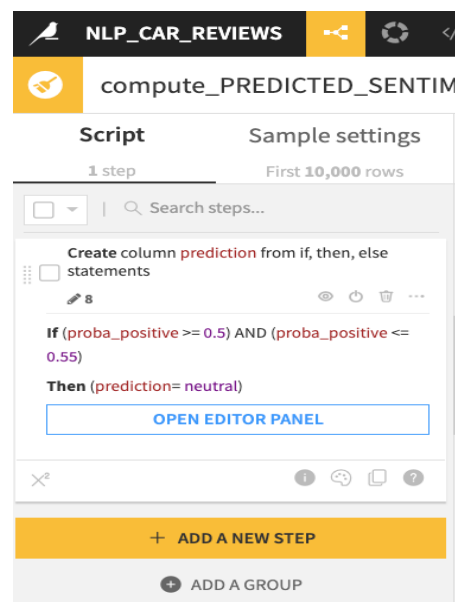
These below image represent whether reviews are seen as positive or negative based on the analysis.



The screenshot shows a data table with 104 rows. The columns include: Year, Model, Review, Normalize, Tokenized, proba_negative, proba_positive, and prediction. The table displays various car reviews and their corresponding sentiment scores and predictions.

Year	Model	Review	Normalize	Tokenized	proba_negative	proba_positive	prediction
2016	Honda	2016 Honda Civic S 1.4: The interior of the car is a v...	2016 honda civic 1 4 interior car veri nice comfort d...	["2016","honda","civic","1","4","interior","car","veri",...	0.17500448292879855	0.8249955170712014	positive
2016	Honda	Our Honda's Pilot is smooth, cushy, and powerful, ...	honda pilot smooth cushi power tranri continu su...	["honda","pilot","smooth","cushi","power","tranni",...	0.46572440379091296	0.534275596209087	positive
2017	Honda	Honda HR-V GH1 1.6: Fell in love with this when I go...	honda hr v gh1 1 6 fell love got 98 one comfort car ...	["honda","hr","v","gh1","1","6","fell","love","got","98",...	0.08798572988603304	0.912014270113967	positive
2017	Honda	The Honda's angular interior is architectural, only I...	honda angular interior architectur frustrat infotain ...	["honda","angular","interior","architecture","frustrat",...	0.1451740862896923	0.8548259137103077	positive
2018	Honda	Honda CR-V: Water pump seized, left us stranded o...	honda cr v water pump seiz left us strand highway ...	["honda","cr","v","water","pump","seiz","left","us","st...",...	0.5895311846313034	0.41046881536869656	negative
2018	Honda	Honda's engineers went to work upgrading the sus...	honda engin went work upgrad suspens steer chas...	["honda","engin","went","work","upgrad","suspens",...	0.7239802471170591	0.2760197528829409	negative
2019	Honda	The 2019 Honda Insight marks the return of a dedic...	2019 honda insight mark return dedic hybrid vehicl...	["2019","honda","insight","mark","return","dedic","h...",...	0.4402841539211526	0.5597158460788474	positive
2016	Hyundai	Hyundai XG350: Very reliable up to ~130k miles, the...	hyundai xg350 veri reliabl 130k mile start fall apart ...	["hyundai","xg350","veri","reliabl","130k","mile","sta...",...	0.2808926790326898	0.7191073209673102	positive
2016	Hyundai	Enter the 2016 Hyundai Elantra Value Edition that o...	enter 2016 hyundai elantra valu edit offer enough g...	["enter","2016","hyundai","elantra","valu","edit","off...",...	0.8010079001811421	0.19899209981885785	negative
2017	Hyundai	2017 Hyundai Elantra GLS: Only beware of dramati...	2017 hyundai elantra gls bewar dramt drop price ...	["2017","hyundai","elantra","gls","bewar","dramat","...",...	0.3137091920428794	0.6862908079571206	positive
2017	Hyundai	Hyundai Elantra : A Fun Little Gem Hiding in Plain S...	hyundai elantra fun littl gem hide plain sight 2017 e...	["hyundai","elantra","fun","littl","gem","hide","plain",...	0.3780869661490598	0.6219130338509402	positive
2018	Hyundai	2018 Hyundai Elantra SEL 2.0L : Overall, I'm very im...	2018 hyundai elantra sel 2 0l overall ver impress sat...	["2018","hyundai","elantra","sel","2","0l","overall","ve...",...	0.28772339670679814	0.7122766032932019	positive
2018	Hyundai	New for the 2018 model year, the Hyundai Elantra ...	new 2018 model year hyundai elantra gt join grow ...	["new","2018","model","year","hyundai","elantra","g...",...	0.23154681603635674	0.7684531839636433	positive
2019	Hyundai	Hyundai Santa Fe Sport Premium : Very well screw...	hyundai santa fe sport premium veri well screw ver...	["hyundai","santa","fe","sport","premium","veri","we...",...	0.1665589564038807	0.8334410435961119	positive
2016	GMC	2016 GMC Terrain Denali: More style, luxury, and fe...	2016 gmc terrain denali style luxuri featur top mod...	["2016","gmc","terrain","denali","style","luxuri","feat...",...	0.41195283892553836	0.5880471610744616	positive
2016	GMC	It's a great, reliable car that I feel safe driving. Mine...	great reliabl car feel safe drive need bodi work got...	["great","reliabl","car","feel","safe","drive","need","b...",...	0.2680396821425456	0.7319603178574544	positive
2017	GMC	The 2017 GMC Acadia rides on a 112.5-inch-long wh...	2017 gmc acadia ride 112 5 inch long wheelbas 193 ...	["2017","gmc","acadia","ride","112","5","inch","long",..."	0.48015957865902803	0.519840421340972	positive
2017	GMC	2017 GMC Acadia : All in all, GMC has met its target ...	2017 gmc acadia gmc met target creat smaller nim...	["2017","gmc","acadia","gmc","met","target","creat",..."	0.44092071709086644	0.5590792829091336	positive
2018	GMC	Key upgrade is the 2018 Yukon Denali becomes one...	key upgrad 2018 yukon denali becom one first suv...	["key","upgrad","2018","yukon","denali","becom","one",..."	0.6723633465229883	0.32763665347701068	negative

Added another condition in prepare recipe that if the given polarity is between 0.5 and 0.55 then the sentiment is Neutral



The screenshot shows a script editor with a single step. The script is as follows:

```
Create column prediction from if, then, else statements

If (proba_positive >= 0.5) AND (proba_positive <= 0.55)
Then (prediction = neutral)

OPEN EDITOR PANEL
```

Below the script, there are buttons for "ADD A NEW STEP" and "ADD A GROUP".

Predicted Sentiments :

The below image displays the final predicted sentiments for each row, corresponding to the customer reviews analyzed.

	Model	Review	Normalize	Tokenized	proba_negative	proba_positive	prediction
	string	string	string	string	double	double	string
	Text	Natural lang.	Natural lang.	Bag of words	Decimal	Decimal	Text
016	Honda	2016 Honda Civic S 1.4: The interior of the car is a v...	2016 honda civic 1 4 interior car veri nice comfort d...	["2016","honda","civic","1","4","interior","car","veri"...	0.17500448292879855	0.8249955170712014	positive
016	Honda	Our Honda's Pilot is smooth, cushy, and powerful, ...	honda pilot smooth cushi power tranni continu su...	["honda","pilot","smooth","cushi","power","tranni",...]	0.46572440379091296	0.534275596209087	neutral
017	Honda	Honda HR V GH1 1.6: Fell in love with this when I go...	honda hr v gh1 1 6 fell love got 98 one comfort car ...	["honda","hr","v","gh1","1","6","fell","love","got","98"...	0.08798572988603304	0.912014270113967	positive
017	Honda	The Honda's angular interior is architectural, only l...	honda angular interior architectur frustrat infotain ...	["honda","angular","interior","architectur","frustrat"...	0.1451740862896923	0.8548259137103077	positive
018	Honda	Honda CR-V : Water pump seized, left us stranded o...	honda cr v water pump seiz left us strand highway ...	["honda","cr","v","water","pump","seiz","left","us","st..."	0.5895311846313034	0.41046881536869656	negative
018	Honda	Honda's engineers went to work upgrading the sus...	honda engin went work upgrad suspens steer chas...	["honda","engin","went","work","upgrad","suspens",...]	0.7239802471170591	0.2760197528829409	negative
019	Honda	The 2019 Honda Insight marks the return of a dedic...	2019 honda insight mark return dedic hybrid vehicl...	["2019","honda","insight","mark","return","dedic","h..."	0.4402841539211526	0.5597158460788474	positive
016	Hyundai	Hyundai XG350: Very reliable up to ~130k miles, the...	hyundai xg350 veri reliabl 130k mile start fail apart...	["hyundai","xg350","veri","reliabl","130k","mile","sta..."	0.2808926790326898	0.7191073209673102	positive
016	Hyundai	Enter the 2016 Hyundai Elantra Value Edition that o...	enter 2016 hyundai elantra valu edit offer enough g...	["enter","2016","hyundai","elantra","valu","edit","off..."	0.8010079001811421	0.19899209981885785	negative
017	Hyundai	2017 Hyundai Elantra GLS: Only beware of dramati...	2017 hyundai elantra gls bewar dramat drop price ...	["2017","hyundai","elantra","gls","bewar","dramat",...]	0.3137091920428794	0.6862908079571206	positive
017	Hyundai	Hyundai Elantra : A Fun Little Gem Hiding in Plain S...	hyundai elantra fun littl gem hide plain sight 2017 e...	["hyundai","elantra","fun","littl","gem","hide","plain"...]	0.3780869661490598	0.6219130338509402	positive
018	Hyundai	2018 Hyundai Elantra SEL 2.0L: Overall, I'm very im...	2018 hyundai elantra sel 2 0l overall impress sat...	["2018","hyundai","elantra","sel","2","0l","overall","ve..."	0.28772339670679814	0.7122766032932019	positive
018	Hyundai	New for the 2018 model year, the Hyundai Elantra ...	new 2018 model year hyundai elantra gt join grow ...	["new","2018","model","year","hyundai","elantra","g..."	0.23154681603635674	0.7684531839636433	positive
019	Hyundai	Hyundai Santa Fe Sport Premium : Very well screw...	hyundai santa fe sport premium veri well screw ver...	["hyundai","santa","fe","sport","premium","veri","we..."	0.16655895640388807	0.8334410435961119	positive
016	GMC	2016 GMC Terrain Denali: More style, luxury, and fe...	2016 gmc terrain denali style luxuri featur top mod...	["2016","gmc","terrain","denali","style","luxuri","feat..."	0.41195283892553836	0.5880471610744616	positive
016	GMC	It's a great, reliable car that I feel safe driving. Mine ...	great reliabl car feel safe drive need bodi work got ...	["great","reliabl","car","feel","safe","drive","need","b..."	0.2680396821425456	0.7319603178574544	positive
017	GMC	The 2017 GMC Acadia rides on a 112.5-inch-long wh...	2017 gmc acadia ride 112 5 inch long wheelbas 193 ...	["2017","gmc","acadia","ride","112","5","inch","long",...]	0.48015957865902803	0.519840421340972	neutral
017	GMC	2017 GMC Acadia : All in all, GMC has met its target ...	2017 gmc acadia gmc met target creat smaller nim...	["2017","gmc","acadia","gmc","met","target","creat",...]	0.44092071709086644	0.5590792829091336	positive