

Logistic Regression

Karol Orozco & Charles Hanks

text feature engineering:

```
wine_words <- function(df, j = 1000, stem=F){
  library(tidytext)
  library(SnowballC)
  data(stop_words)
  words <- df %>%
    unnest_tokens(word, description) %>%
    anti_join(stop_words) %>% # get rid of stop words
    filter(!(word %in% c("wine", "pinot", "vineyard", "price", "points")))

  if(stem){
    words <- words %>%
      mutate(word = wordStem(word))
  }

  words <- words %>%
    count(id, word) %>%
    group_by(id) %>%
    mutate(exists = (n>0)) %>%
    ungroup %>%
    group_by(word) %>%
    mutate(total = sum(n)) %>%
    filter(total > j) %>%
    pivot_wider(id_cols = id, names_from = word, values_from = exists, values_fill = list(
    right_join(select(df, id, province)) %>%
    mutate(across(-province, ~replace_na(.x, F)))
  )
  wino <- wine_words(wine, j=400, stem=F)
```

Joining with `by = join_by(word)`

Joining with `by = join_by(id)`

bringing back numerical features from original dataset to wino:

```
wino = wino %>% left_join(select(wine, id, price, points, year), by = "id")
```

Numerical feature engineering:

```
#center and scale points:
wino = wino %>% select(points) %>% preProcess(method = c("center", "scale")) %>% predict(w

#year as factor, logprice:
wino = wino %>% mutate(year_f = as.factor(year),
                        lprice = log(price))

#binning year and price:
wino = wino %>%
  mutate(price_f = case_when(
    price < 16 ~ "low",
    price >= 16 & price < 41 ~ "med",
    price >= 41 ~ "high"
  ),
  year_f = case_when(
    year < 2005 ~ "old",
    year >= 2005 & year < 2011 ~ "recent",
    year >= 2011 ~ "current"
  ))
wino = wino %>% dplyr::select(-price)
#difference of wine's lprice from total average lprice
wino = wino %>% mutate(diff_from_avg_lprice = mean(lprice) - lprice)
wino = wino %>% mutate(cost_per_point = lprice/points)
wino = wino %>% select(-id)
wino = wino %>% select(-diff_from_avg_lprice)

head(wino)
```

A tibble: 6 x 82

	bottling	earthy	herbal	berry	chocolate	drink	herb	oak	tart	aromas	bodied
	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>	<lgl>
1	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
2	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE
3	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE

```

4 FALSE    FALSE TRUE    FALSE FALSE    FALSE TRUE  FALSE FALSE FALSE  FALSE
5 FALSE    TRUE  FALSE TRUE    FALSE    FALSE FALSE FALSE FALSE FALSE  FALSE
6 FALSE    FALSE FALSE FALSE FALSE    TRUE  FALSE FALSE FALSE FALSE  FALSE
# ... with 71 more variables: earth <lgl>, forest <lgl>, offers <lgl>,
#   raspberry <lgl>, smooth <lgl>, spice <lgl>, texture <lgl>, finish <lgl>,
#   flavor <lgl>, fruit <lgl>, notes <lgl>, sweet <lgl>, touch <lgl>,
#   flavors <lgl>, tannins <lgl>, fruity <lgl>, strawberry <lgl>,
#   cranberry <lgl>, dark <lgl>, palate <lgl>, acidity <lgl>, black <lgl>,
#   cherry <lgl>, cola <lgl>, dried <lgl>, nose <lgl>, soft <lgl>, juicy <lgl>,
#   ripe <lgl>, light <lgl>, spicy <lgl>, red <lgl>, age <lgl>, bit <lgl>, ...

```

Split the data

```

set.seed(1000)
wine_index <- createDataPartition(wino$province, p = 0.80, list = FALSE)

train <- wino[ wine_index, ]
test  <- wino[-wine_index, ]

table(train$province)

```

Burgundy	California	Casablanca_Valley	Marlborough
955	3168	105	184
New_York	Oregon		
105	2190		

```

# Fit the mode

control <- trainControl(method="cv",
                        number=10,
                        savePredictions="all",
                        classProbs=TRUE)

model <- nnet::multinom(province ~.,
                        data = train,
                        trControl=control)

```

```

# weights:  510 (420 variable)

```

```

initial value 12017.330760
iter 10 value 8584.678063
iter 20 value 3590.850590
iter 30 value 2901.827100
iter 40 value 2452.435560
iter 50 value 2219.885203
iter 60 value 2083.635497
iter 70 value 1970.867905
iter 80 value 1917.319694
iter 90 value 1857.675672
iter 100 value 1832.662495
final value 1832.662495
stopped after 100 iterations

```

```

# Summarize the model
summary(model)

```

Call:

```
nnet::multinom(formula = province ~ ., data = train, trControl = control)
```

Coefficients:

	(Intercept)	bottlingTRUE	earthyTRUE	herbalTRUE	berryTRUE
California	-2.77500648	2.5209736	2.2453710	1.5339226	-1.09657930
Casablanca_Valley	-0.10157961	0.4768997	1.8774559	4.0175335	1.25237824
Marlborough	0.01963430	1.9118278	-0.1145254	1.6625478	-1.37893903
New_York	-0.06825329	-0.8768964	1.3880775	0.9986137	0.77942113
Oregon	1.79051669	1.5560276	0.7711434	1.0970615	-0.04907631
	chocolateTRUE	drinkTRUE	herbTRUE	oakTRUE	tartTRUE
California	-0.6643259	-3.3716573	3.436038	3.079318	3.100881
Casablanca_Valley	0.8297156	-0.9351729	1.301126	3.813594	2.411478
Marlborough	0.1048103	1.4577371	3.358910	2.875729	3.757737
New_York	-0.2232008	-2.5937527	4.883826	2.600138	4.099524
Oregon	2.1629387	-1.9855106	3.787791	2.242853	3.794446
	aromasTRUE	bodiedTRUE	earthTRUE	forestTRUE	offersTRUE
California	3.987299	2.9749092	4.066725	4.679266	1.1833147
Casablanca_Valley	5.896640	2.0703769	3.135205	2.444916	-1.0393156
Marlborough	2.873201	3.6866901	4.434232	4.569610	0.5302051
New_York	2.931192	3.5573119	3.852640	3.161319	0.7286386
Oregon	1.329239	0.8036162	3.946574	2.632343	0.7777891
	raspberryTRUE	smoothTRUE	spiceTRUE	textureTRUE	finishTRUE
California	0.6062337	1.1536197	-0.1726341	0.7045566	1.409196

Casablanca_Valley	2.8150701	2.0737349	0.8693627	-3.2567182	3.796586	
Marlborough	-1.5329222	0.6252891	-2.3327408	0.7286398	2.806587	
New_York	0.8630589	-2.6231495	-0.3213235	-4.7358893	2.346425	
Oregon	0.5319400	1.3031381	-1.5063145	-1.8071877	1.775728	
	flavorTRUE	fruitTRUE	notesTRUE	sweetTRUE	touchTRUE	
California	0.5129564	-0.7700230	2.332698	0.4876928	-1.845129	
Casablanca_Valley	1.4119770	-0.7809294	4.345764	1.4633271	-1.584681	
Marlborough	-0.8182311	0.7906989	4.356205	-2.7113104	-1.850882	
New_York	1.7685619	-1.9319092	4.802967	0.2811852	-2.449665	
Oregon	0.4023136	1.0811985	2.690080	0.8509886	-1.419243	
	flavorsTRUE	tanninsTRUE	fruityTRUE	strawberryTRUE		
California	-0.27103187	-2.2722930	-1.853853	-0.679039		
Casablanca_Valley	2.51289914	-3.5636578	-1.250639	-1.876021		
Marlborough	-0.69649195	-1.0574806	-4.972996	-2.158214		
New_York	-0.03926085	0.8133222	-2.255440	-1.159168		
Oregon	0.88531711	-1.8448261	-1.352891	-1.086874		
	cranberryTRUE	darkTRUE	palateTRUE	acidityTRUE	blackTRUE	
California	2.6986905	-0.26088270	1.8395931	-1.856400	0.22377633	
Casablanca_Valley	0.4285337	-0.09059593	2.9407091	-1.553167	-0.02662684	
Marlborough	1.9133996	-1.12332313	1.1804939	-5.476348	0.19441739	
New_York	2.5996938	-4.21790610	3.5186774	-1.058879	1.08744301	
Oregon	1.9567369	-2.25848909	0.6028781	-3.601053	-0.60808725	
	cherryTRUE	colaTRUE	driedTRUE	noseTRUE	softTRUE	juicyTRUE
California	2.026573	3.877221	3.710345	3.3457571	-1.526021	-1.6983668
Casablanca_Valley	2.022925	3.037918	4.227264	3.6964722	-2.943853	-0.8147643
Marlborough	2.440956	3.942322	3.247060	0.9568684	-2.997071	-8.9815429
New_York	3.750786	2.370356	4.894553	2.1552595	-1.300078	-1.3778476
Oregon	1.797066	4.570044	2.430325	0.3120050	-3.006307	-2.1172969
	ripeTRUE	lightTRUE	spicyTRUE	redTRUE	ageTRUE	
California	-0.310123106	-0.5044199	2.1268473	-2.09406904	-1.8282947	
Casablanca_Valley	0.009579719	-0.9561360	2.8327447	-0.78570852	-1.6148570	
Marlborough	-1.805042178	-1.4096910	-2.4202588	-4.02642933	-1.5628434	
New_York	0.442156692	-1.3891852	-0.8157115	-0.06465313	-0.9315412	
Oregon	-1.132798172	-0.3082791	1.7785449	-3.38248702	-0.3135615	
	bitTRUE	tightTRUE	cherriesTRUE	coreTRUE	fruitsTRUE	
California	2.821779	-1.648567	1.8510042	-0.8577396	-2.57366773	
Casablanca_Valley	3.078756	-2.677989	-0.1840197	-2.1164158	-0.67466774	
Marlborough	3.529358	-2.983731	4.0697903	-2.8913075	-2.48596178	
New_York	4.158226	-3.183529	2.2038128	-1.7483764	-3.25816470	
Oregon	4.046066	-1.095682	2.3548699	-0.2130100	0.04856493	
	richTRUE	agingTRUE	brightTRUE	characterTRUE		
California	-1.2669757	-3.776425	1.0289868	-2.024692		
Casablanca_Valley	-1.2208220	-2.771589	1.1068174	-3.395049		

Marlborough	-1.7660246	-4.255738	1.6195382	-5.174097		
New_York	-0.4932591	-2.236244	2.7419775	-1.938130		
Oregon	-2.4651431	-1.357490	0.1593359	-2.613282		
	concentratedTRUE	vintageTRUE	complexTRUE	estateTRUE	teaTRUE	
California	0.9467508	-0.4821498	-0.9185983	1.083988	3.8850754	
Casablanca_Valley	-2.5786493	0.1950360	-2.4965566	1.515273	3.9280504	
Marlborough	-0.5702413	-1.4622083	1.3451000	-4.128839	0.4366901	
New_York	2.7822444	-2.1289379	-0.1315482	-2.455003	0.3844416	
Oregon	-0.1958550	0.2326041	-0.3847067	1.427878	2.9612566	
	wildTRUE	firmTRUE	noirTRUE	mediumTRUE	structureTRUE	
California	0.24021942	-0.3389080	1.0344024	1.45088421	-2.2425752	
Casablanca_Valley	-1.28548790	-2.3655243	1.9480954	0.55061391	-0.9459815	
Marlborough	-4.75889564	-0.5749532	2.4172721	2.52742259	-4.6092447	
New_York	-0.65067923	-2.0268298	4.0831641	0.06512319	-2.6658477	
Oregon	-0.06725654	-1.5791229	0.8699755	0.87686594	-3.8460994	
	cloveTRUE	timeTRUE	freshTRUE	balancedTRUE	structuredTRUE	
California	4.822613	-1.007236	-0.7086727852	0.2001176	-2.031786	
Casablanca_Valley	3.221613	-2.473415	0.0006037306	1.2844956	-3.725305	
Marlborough	3.898832	-1.303105	-1.3078335889	-3.3291800	-1.825072	
New_York	2.807813	-2.692847	-0.5539863349	-0.1136392	-4.128570	
Oregon	3.061582	-1.516108	-0.7289361006	0.1074632	-3.399736	
	orangeTRUE	plumTRUE	pomegranateTRUE	cinnamonTRUE	savoryTRUE	
California	4.2612356	-1.0193761	7.969549	2.2873118	5.831156	
Casablanca_Valley	3.2123455	1.8624871	5.737819	1.3164538	4.714932	
Marlborough	0.7151617	-1.7072324	7.779114	0.8622363	7.189825	
New_York	1.5796032	-0.1125171	5.130179	1.3998277	7.563856	
Oregon	3.7475928	-0.7895521	6.202471	1.5734575	3.121859	
	pepperTRUE	roseTRUE	points	year	year_fold	
California	5.476676	2.752658	0.8112382	0.011961956	-7.023546	
Casablanca_Valley	5.341680	-1.457274	-0.5157338	0.007071878	-4.759791	
Marlborough	4.454945	1.597131	0.7672799	0.012071212	-3.690217	
New_York	2.443789	2.168767	0.1329819	0.007962507	-5.828934	
Oregon	5.232729	1.851036	1.0516143	0.010926753	-3.338695	
	year_frecent	lprice	price_flow	price_fmcd	cost_per_point	
California	-3.68228628	-4.119757	-3.425767	-2.87693973	0.0004341442	
Casablanca_Valley	-0.14204301	-5.552746	-2.805696	-3.11207422	0.0046429021	
Marlborough	0.02246605	-6.482572	-3.648329	-1.12006686	-0.0016645341	
New_York	-1.38459326	-5.799768	-1.342581	-0.03329879	0.0008014719	
Oregon	-1.52487526	-4.521092	-4.558390	-2.78805259	0.0005621888	

Std. Errors:

	(Intercept)	bottlingTRUE	earthyTRUE	herbalTRUE	berryTRUE	
California	1.719415e-04	0.0394793232	0.087907905	0.038084025	0.053270448	

Casablanca_Valley	2.302220e-05	0.0016258667	0.001312199	0.002185899	0.002052613
Marlborough	6.553365e-05	0.0018666714	0.003683545	0.003740238	0.002947693
New_York	3.443324e-05	0.0004793587	0.002678153	0.001917216	0.004522157
Oregon	2.172732e-04	0.0377067633	0.088836386	0.040515392	0.059391266
	chocolateTRUE	drinkTRUE	herbTRUE	oakTRUE	
California	0.0441731689	0.071973013	0.0710594337	0.069247510	
Casablanca_Valley	0.0018219386	0.002032301	0.0006492517	0.002437634	
Marlborough	0.0041339151	0.017070348	0.0032010114	0.005678034	
New_York	0.0006576456	0.001791696	0.0027441881	0.001054560	
Oregon	0.0508636092	0.093763747	0.0715165408	0.070269284	
	tartTRUE	aromasTRUE	bodiedTRUE	earthTRUE	forestTRUE
California	0.081963724	0.086589470	0.079866042	0.0857243274	0.0081274089
Casablanca_Valley	0.001070397	0.004674932	0.002569986	0.0007230178	0.0002930786
Marlborough	0.008310574	0.009717525	0.016455318	0.0053041067	0.0019518867
New_York	0.004431897	0.007071638	0.004681259	0.0024858586	0.0007759841
Oregon	0.083732732	0.073910405	0.061945572	0.0848192269	0.0076321932
	offersTRUE	raspberryTRUE	smoothTRUE	spiceTRUE	
California	0.087419824	0.070821830	0.040343387	0.087089761	
Casablanca_Valley	0.001333962	0.005349329	0.001239403	0.004757900	
Marlborough	0.002830329	0.001521426	0.002151977	0.003998272	
New_York	0.002418715	0.003504430	0.001062236	0.006086569	
Oregon	0.088639640	0.070207258	0.042306665	0.080974248	
	textureTRUE	finishTRUE	flavorTRUE	fruitTRUE	notesTRUE
California	0.0780325598	0.060873188	0.088450345	0.053212183	0.088658199
Casablanca_Valley	0.0016971488	0.002797241	0.001958817	0.002250438	0.002177757
Marlborough	0.0123597244	0.006799026	0.003058842	0.013005603	0.005749720
New_York	0.0001145886	0.004293504	0.005283677	0.003155229	0.003622139
Oregon	0.0656688277	0.062040702	0.088878121	0.055377361	0.092009440
	sweetTRUE	touchTRUE	flavorsTRUE	tanninsTRUE	fruityTRUE
California	0.017990082	0.080793755	0.054084109	0.066680302	0.016932742
Casablanca_Valley	0.001467188	0.002540427	0.002440261	0.001017936	0.001434643
Marlborough	0.001037824	0.002671071	0.010504045	0.007355396	0.001233637
New_York	0.001999442	0.001857465	0.006740494	0.007405988	0.001351809
Oregon	0.017750115	0.083589097	0.060303530	0.073467677	0.018582646
	strawberryTRUE	cranberryTRUE	darkTRUE	palateTRUE	
California	0.075609251	0.082653198	0.074432324	0.070408218	
Casablanca_Valley	0.002439287	0.001271257	0.003330679	0.005128075	
Marlborough	0.001938390	0.003233459	0.001960605	0.005756635	
New_York	0.003218366	0.003206557	0.001155100	0.006199995	
Oregon	0.075990974	0.081931207	0.073680606	0.068055830	
	acidityTRUE	blackTRUE	cherryTRUE	colaTRUE	driedTRUE
California	0.076940603	0.064847821	0.054818437	0.0882327689	0.057460834
Casablanca_Valley	0.003959028	0.005408641	0.004403557	0.0007707169	0.001618321

Marlborough	0.001754184	0.014941077	0.007528584	0.0037476283	0.002195061
New_York	0.007955056	0.009697174	0.009352051	0.0009487048	0.005344558
Oregon	0.073980038	0.064329632	0.056231843	0.0887535566	0.053909719
	noseTRUE	softTRUE	juicyTRUE	ripeTRUE	lightTRUE
California	0.072455432	0.0616294481	5.209954e-02	0.078452971	0.071857850
Casablanca_Valley	0.005009358	0.0008009083	1.030901e-03	0.002944626	0.003425548
Marlborough	0.002624505	0.0040482838	8.044305e-06	0.003674383	0.006403373
New_York	0.006345070	0.0027865352	2.515925e-03	0.006297878	0.003266251
Oregon	0.062671018	0.0606514214	5.155912e-02	0.078623059	0.075748857
	spicyTRUE	redTRUE	ageTRUE	bitTRUE	tightTRUE
California	0.0836405368	0.069898393	0.012176591	0.037776528	0.0184629342
Casablanca_Valley	0.0024791069	0.001961892	0.000211958	0.002430723	0.0008802425
Marlborough	0.0010972826	0.002684283	0.002433669	0.004890361	0.0007041964
New_York	0.0003355526	0.004152037	0.001878453	0.002736310	0.0006339256
Oregon	0.0825284598	0.070242729	0.015836031	0.039504487	0.0191626225
	cherriesTRUE	coreTRUE	fruitsTRUE	richTRUE	
California	0.0325663675	0.0217601252	0.0261024394	0.055161886	
Casablanca_Valley	0.0001347566	0.0009772601	0.0013329558	0.002257235	
Marlborough	0.0077947824	0.0004318635	0.0006831784	0.007168181	
New_York	0.0020358032	0.0006688722	0.0004829530	0.004845049	
Oregon	0.0326836227	0.0224875161	0.0336928576	0.051059350	
	agingTRUE	brightTRUE	characterTRUE	concentratedTRUE	
California	0.0069742849	0.052582925	0.017521080	2.757774e-02	
Casablanca_Valley	0.0002993544	0.001799931	0.001236932	5.762273e-05	
Marlborough	0.0008842052	0.001578567	0.000862982	2.693995e-03	
New_York	0.0011600937	0.003013410	0.001123660	2.965611e-03	
Oregon	0.0106321324	0.051286758	0.020000987	2.749381e-02	
	vintageTRUE	complexTRUE	estateTRUE	teaTRUE	
California	0.0316442503	0.0333758737	4.817880e-02	0.0406954217	
Casablanca_Valley	0.0012313132	0.0008333883	1.067598e-03	0.0008163434	
Marlborough	0.0021493085	0.0064585111	1.267738e-03	0.0019724206	
New_York	0.0002213314	0.0016401929	9.191167e-05	0.0005633459	
Oregon	0.0335850153	0.0331954588	4.913375e-02	0.0397744784	
	wildTRUE	firmTRUE	noirTRUE	mediumTRUE	
California	2.576501e-02	0.0210422874	0.079235340	0.057539259	
Casablanca_Valley	2.842943e-04	0.0003710987	0.001808535	0.001193553	
Marlborough	1.930763e-05	0.0035849085	0.010505599	0.014754129	
New_York	1.271847e-03	0.0016958175	0.004616756	0.003867468	
Oregon	2.581474e-02	0.0211023913	0.084152399	0.044554485	
	structureTRUE	cloveTRUE	timeTRUE	freshTRUE	
California	0.011737111	0.0149552783	1.812303e-02	0.091008764	
Casablanca_Valley	0.001035628	0.0007171999	3.477994e-04	0.002197237	
Marlborough	0.001368301	0.0017974986	2.330943e-03	0.002746505	

New_York	0.002725939	0.0003304636	8.377225e-05	0.004851295
Oregon	0.012670282	0.0143281293	1.902628e-02	0.090905478
	balancedTRUE	structuredTRUE	orangeTRUE	plumTRUE
California	0.059505843	0.0087208711	3.008611e-02	0.087330910
Casablanca_Valley	0.001344610	0.0005063211	5.030824e-04	0.006430368
Marlborough	0.000549597	0.0022073885	9.097799e-05	0.002475507
New_York	0.001124548	0.0013622686	3.643664e-04	0.005203934
Oregon	0.061546732	0.0084426662	3.002776e-02	0.084022793
	pomegranateTRUE	cinnamonTRUE	savoryTRUE	pepperTRUE
California	0.0168131429	0.057769922	0.0103968743	0.0325870496
Casablanca_Valley	0.0004547328	0.001370427	0.0006049508	0.0008551444
Marlborough	0.0022832782	0.001967276	0.0056086561	0.0012357472
New_York	0.0010206977	0.002791881	0.0020760534	0.0001293620
Oregon	0.0151833779	0.054275007	0.0073133822	0.0320672258
	roseTRUE	points	year	year_fold year_frecent
California	0.0475323643	0.06656363	1.500173e-04	0.012968790 0.062757704
Casablanca_Valley	0.0004990768	0.02918698	1.260220e-04	0.001815310 0.004902305
Marlborough	0.0019190247	0.10376309	9.139525e-05	0.003856658 0.020208260
New_York	0.0030651546	0.05994255	1.077433e-04	0.001823583 0.008490916
Oregon	0.0465206936	0.06497081	1.489254e-04	0.016984034 0.085130346
	lprice	price_flow	price_fmed	cost_per_point
California	0.07065033	0.042164641	0.063940185	0.0007900074
Casablanca_Valley	0.01124685	0.006453263	0.002997031	0.0013747144
Marlborough	0.02932427	0.007170623	0.013173927	0.0012306636
New_York	0.01673558	0.008638582	0.006079677	0.0013903580
Oregon	0.06867945	0.035336134	0.060005391	0.0007652392

Residual Deviance: 3665.325

AIC: 4505.325

```
# Make predictions
predicted.classes <- model %>% predict(test)
head(predicted.classes)
```

```
[1] Oregon    Oregon    California California California Oregon
6 Levels: Burgundy Casablanca_Valley Marlborough ... Oregon
```

```
# Model accuracy
mean(predicted.classes == test$province)
```

```
[1] 0.8828452
```

```
varImp(model)%>%
  arrange(desc(Overall))
```

	Overall
pomegranateTRUE	32.819130475
savoryTRUE	28.421627534
lprice	26.475934835
year_fold	24.641182561
pepperTRUE	22.949819354
earthTRUE	19.435376728
notesTRUE	18.527713356
driedTRUE	18.509548695
cloveTRUE	17.812451886
colaTRUE	17.797860972
bitTRUE	17.634184899
forestTRUE	17.487454152
tartTRUE	17.164065879
aromasTRUE	17.017570445
herbTRUE	16.767691647
price_flow	15.780762880
characterTRUE	15.145249894
structuredTRUE	15.110468984
juicyTRUE	14.989818627
oakTRUE	14.611632867
agingTRUE	14.397484962
structureTRUE	14.309748544
acidityTRUE	13.545846805
orangeTRUE	13.515938671
bodiedTRUE	13.092904284
finishTRUE	12.134521551
cherryTRUE	12.038306784
softTRUE	11.773329594
fruityTRUE	11.685817634
teaTRUE	11.595514045
tightTRUE	11.589497642
textureTRUE	11.232991436
cherriesTRUE	10.663497001
estateTRUE	10.610979731
noseTRUE	10.466362220
redTRUE	10.353347040
noirTRUE	10.352909503
drinkTRUE	10.343830697

palateTRUE	10.082351630
spicyTRUE	9.974107164
price_fmed	9.930432187
roseTRUE	9.826866097
cranberryTRUE	9.597054458
tanninsTRUE	9.551579722
herbalTRUE	9.309679118
touchTRUE	9.149598987
fruitsTRUE	9.041026868
timeTRUE	8.992711975
darkTRUE	7.951196939
coreTRUE	7.826849298
smoothTRUE	7.778931272
cinnamonTRUE	7.439287158
bottlingTRUE	7.342625204
richTRUE	7.212224540
concentratedTRUE	7.073740763
wildTRUE	7.002538730
strawberryTRUE	6.959315672
firmTRUE	6.885338184
year_frecent	6.756263864
brightTRUE	6.656655756
earthyTRUE	6.396573214
raspberryTRUE	6.349224841
ageTRUE	6.251097823
sweetTRUE	5.794504124
plumTRUE	5.491164823
mediumTRUE	5.470909838
fruitTRUE	5.354759056
complexTRUE	5.276509741
spiceTRUE	5.202375564
balancedTRUE	5.034895537
flavorTRUE	4.914039999
lightTRUE	4.567711289
berryTRUE	4.556394021
vintageTRUE	4.500936083
flavorsTRUE	4.405000928
offersTRUE	4.259263040
chocolateTRUE	3.984991318
ripeTRUE	3.699699866
freshTRUE	3.300032540
points	3.278848071
blackTRUE	2.140350825

year	0.049994306
cost_per_point	0.008105241