

Ananya Bhupathipalli
Xintao Lu
Omkar Mutreja
Luigi Penaloza
Ruonan Wang



Sports Drinks Analysis

Contents

- Overview
- Selection Criteria
- Data Preparation
- Data Analysis

Selection Criteria

- Selected two premium brands and one low price brand.
- Premium brands: All Sport Lemon Lime and All Sport Cherry Slam
- Low priced brand: Powerade Tidal Burst



Selection Criteria

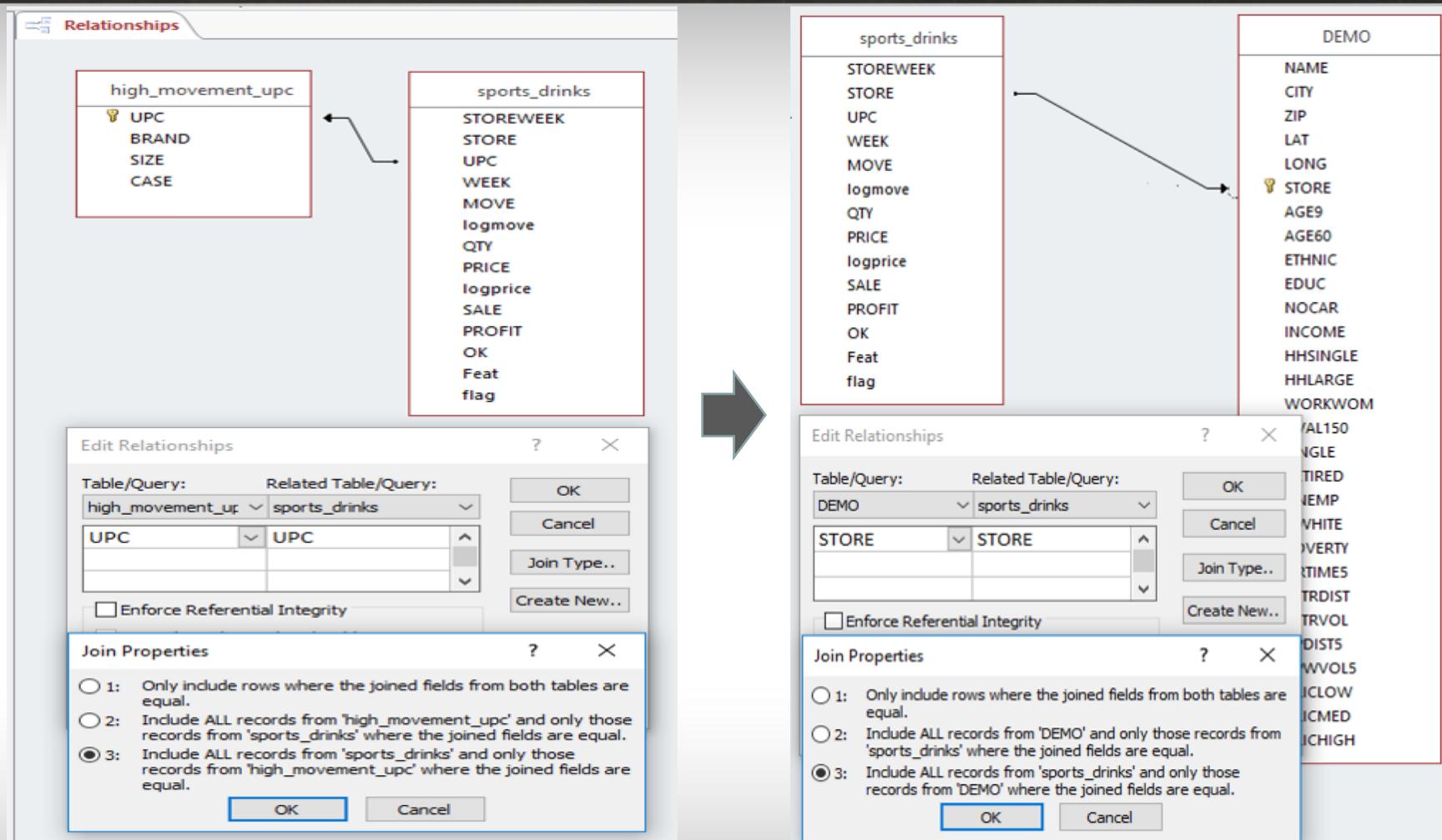
- Sports drinks high movement UPC

UPC	BRAND	SIZE	CASE
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12
1200000735	ALL SPORT LEMON LIME	32 OZ	12
1200000757	ALL SPORT FRUIT PUNC	32 OZ	12
5200003805	GATORADE FRUIT PUNCH	32 OZ	12
5200003925	GATORADE LEMON/LIME	32 OZ	12
5200003940	GATORADE ORANGE DRIN	32 OZ	12
5200032810	GATORADE SPRTS BTL	20 OZ	24
5200032814	GATORADE SPTS BTL F	20 OZ	24
5200032841	GTRADE SPTS BTL CL	20 OZ	24
5200032842	GATORADE SPSTS BTL	20 OZ	24
5200032873	GATORADE WATERMELON-	32 OZ	12
5200033820	GATORADE COOL BLUE R	32 OZ	12
5200033830	GATORADE FRUIT PUNCH	64 OZ	8
5200033831	GATORADE ORANGE	64 OZ	8
5200033832	GATORADE LEMON LIME	64 OZ	8
5200033833	GATORADE LEMONADE	64 OZ	8
5200033934	GATORADE LEMON ICE	32 OZ	12
4900001923	POWERADE FRUIT PUNCH	32 OZ	12
4900002314	POWERADE MOUNTAIN BL	32 OZ	12
4900002450	POWERADE TIDAL BURST	32 OZ	12

Data Preparation

- Establish relationships between sports drinks all data, sports drinks high movement UPC, and Demo small
- Set criteria demographics “**Is Not Null**”
- Create table “new subset” with all needed data using Query design command in Access
- Create a new factor “SEASON” using IIF expression”
- REM: [WEEK] Mod 52
- SEASON: IIF([REM]>10 And [REM]<24, “WINTER”, IIF([REM]>23 And [REM]<37, “SPRING”, IIF([REM]>36 And [REM]<50, “SUMMER”, “FALL”)))

Data Preparation



Data Preparation

Navigation Pane

new subset

UPC	BRAND	SIZE	CASE	STOREWEEK	STORE	WEEK	REM	SEASON	MOVE	logmove	QTY	PRICE	logprice
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2363	2	363	51	FALL	5	1.6094379124341	1	1.31	0.27002713721
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2364	2	364	0	FALL	3	1.09861228866811	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2365	2	365	1	FALL	8	2.07944154167984	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2366	2	366	2	FALL	6	1.79175946922805	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2367	2	367	3	FALL	6	1.79175946922805	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2368	2	368	4	FALL	6	1.79175946922805	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	2372	2	372	8	FALL	2	0.693147180559945	1	1.39	0.3293037471
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	5352	5	352	40	SUMMER	1	0	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8352	8	352	40	SUMMER	6	1.79175946922805	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8353	8	353	41	SUMMER	4	1.38629436111989	1	1.01	9.95033085316809E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8354	8	354	42	SUMMER	2	0.693147180559945	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8355	8	355	43	SUMMER	12	2.484906649788	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8362	8	362	50	FALL	47	3.85014760171006	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8363	8	363	51	FALL	8	2.07944154167984	1	1.05	0.048790164169
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8364	8	364	0	FALL	6	1.79175946922805	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8365	8	365	1	FALL	12	2.484906649788	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8366	8	366	2	FALL	15	2.70805020110221	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8367	8	367	3	FALL	12	2.484906649788	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8368	8	368	4	FALL	7	1.94591014905531	1	0.89	-0.116533816255
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8369	8	369	5	FALL	3	1.09861228866811	1	1.09	8.61776962410524E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8370	8	370	6	FALL	3	1.09861228866811	1	1.09	8.61776962410524E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8371	8	371	7	FALL	1	0	1	1.09	8.61776962410524E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	8373	8	373	9	FALL	12	2.484906649788	1	1.09	8.61776962410524E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9347	9	347	35	SPRING	1	0	1	1.29	0.254642218373
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9352	9	352	40	SUMMER	1	0	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9354	9	354	42	SUMMER	2	0.693147180559945	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9357	9	357	45	SUMMER	2	0.693147180559945	1	1.29	0.254642218373
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9359	9	359	47	SUMMER	1	0	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9361	9	361	49	SUMMER	1	0	1	0.99	-1.00503358535015E
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9363	9	363	51	FALL	2	0.693147180559945	1	1.14	0.131028262406
1200000315	ALLSPORT CHERRY SLAM	32 OZ	12	9365	9	365	1	FALL	4	1.38629436111989	1	0.89	-0.116533816255

Data Analysis

- How does the demand for a brand depend on price?
What is the price elasticity of demand of a brand?

Call:

```
lm(formula = logmove ~ AGE9 + AGE60 + BRAND + EDUC + ETHNIC +
  Feat + HHLARGE + HH SINGLE + HVALL50 + INCOME + logprice +
  NOCAR + NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE +
  STOREWEEK + UNEMP + WORKWOM + BRAND * logprice, data = SportsDrinks)
```

NWHITE	-3.0066939605	0.3602895106	-8.345	< 2e-16	***	
POVERTY	3.9490646447	1.4734660796	2.680	0.007374	**	
REM	0.0112568075	0.0008212779	13.706	< 2e-16	***	
RETIRED	-4.2512678468	1.4722030047	-2.888	0.003891	**	
SEASON[T.SPRING]	-0.5485208377	0.0327033350	-16.773	< 2e-16	***	
SEASON[T.SUMMER]	-0.1099196504	0.0327954755	-3.352	0.000807	***	
SEASON[T.WINTER]	-0.3511093478	0.0329293977	-10.662	< 2e-16	***	
SINGLE	0.4970578117	0.8077827948	0.615	0.538350		
STOREWEEK	0.0000041258	0.0000003287	12.552	< 2e-16	***	
UNEMP	13.9747482465	2.2795945965	6.130	9.17e-10	***	
WORKWOM	5.4865128497	1.1983362673	4.578	4.75e-06	***	
BRAND[T.ALLSPORT CHERRY SLAM]:logprice	-0.2891685803	0.1810658952	-1.597	0.110296		
BRAND[T.POWERADE TIDAL BURST]:logprice	0.1843604778	0.1963734749	0.939	0.347847		

Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.' 1

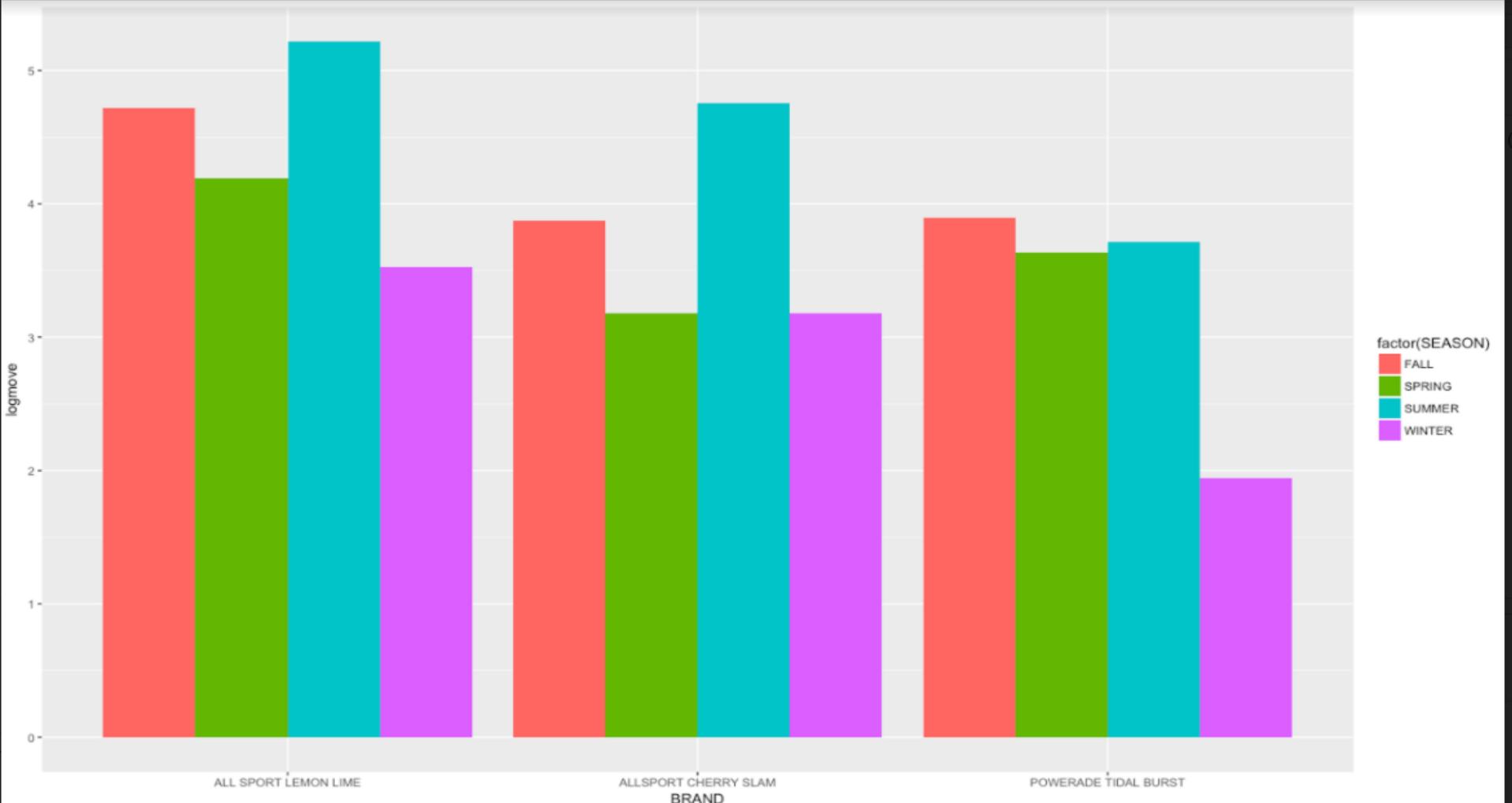
Residual standard error: 0.8172 on 8322 degrees of freedom
Multiple R-squared: 0.3158, Adjusted R-squared: 0.3136
F-statistic: 147.7 on 26 and 8322 DF, p-value: < 2.2e-16

**ALL SPORT LEMON
LIME: -1.9679988575**

**ALLSPORT CHERRY
SLAM: -1.9679988575 -
0.2891685803 =
-1.6788**

**POWERADE TIDAL
BURST: -1.9679988575
+ 0.1843604778 = -
1.7836**

Data Analysis



Data Analysis

- How does demand depend on whether the product is on sale (Feat =1)?

```
Call:  
lm(formula = logmove ~ AGE9 + AGE60 + BRAND + EDUC + ETHNIC +  
    Feat + HHLARGE + HH SINGLE + HVAL150 + INCOME + logprice +  
    NOCAR + NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE +  
    +STOREWEEK + UNEMP + WORKWOM + INCOME + logprice + NOCAR +  
    NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE + STOREWEEK +  
    UNEMP + WORKWOM + BRAND * Feat, data = SportsDrinks)  
  
---  
NOCAR                      1.3682202340   0.3686223366   3.712  0.000207 ***  
NWHITE                     -2.9990492638   0.3601261975  -8.328 < 2e-16 ***  
POVERTY                    3.9248749278   1.4727328892   2.665  0.007713 **  
REM                        0.0109374237   0.0008148545  13.423 < 2e-16 ***  
RETIRED                   -4.2387109938   1.4713550006  -2.881  0.003977 **  
SEASON[T.SPRING]           -0.5454425570   0.0326981122  -16.681 < 2e-16 ***  
SEASON[T.SUMMER]            -0.1003607175   0.0326516708  -3.074  0.002121 **  
SEASON[T.WINTER]            -0.3645311652   0.0329544873  -11.062 < 2e-16 ***  
SINGLE                     0.5251084984   0.8074069217   0.650  0.515475  
STOREWEEK                  0.00000041328   0.00000003286  12.579 < 2e-16 ***  
UNEMP                      13.9515194575   2.2781765170   6.124  9.54e-10 ***  
WORKWOM                    5.4854284624   1.1975824063   4.580  4.71e-06 ***  
BRAND[T.ALLSPORT CHERRY SLAM]:Feat 0.1803015677   0.0548826450   3.285  0.001023 **  
BRAND[T.POWERADE TIDAL BURST]:Feat 0.0958139509   0.0488175723   1.963  0.049715 *  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.8168 on 8322 degrees of freedom  
Multiple R-squared:  0.3164, Adjusted R-squared:  0.3143  
F-statistic: 148.2 on 26 and 8322 DF,  p-value: < 2.2e-16
```

ALL SPORT LEMON LIME:
0.1809583378

ALLSPORT CHERRY SLAM:
0.1803015677 +
0.1809583378 =
0.3618

POWERADE TIDAL BURST:
0.0958139509 +
0.1809583378 =
0.2767

Data Analysis

- What demographic factors affect demand?

```
Call:
lm(formula = logmove ~ AGE9 + AGE60 + BRAND + EDUC + ETHNIC +
    Feat + HHLarge + HHSINGLE + HVAL150 + INCOME + logprice +
    NOCAR + NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE +
    +STOREWEEK + UNEMP + WORKWOM + INCOME + logprice + NOCAR +
    NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE + STOREWEEK +
    UNEMP + WORKWOM + BRAND * Feat, data = SportsDrinks)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.93640	-0.49447	0.07801	0.55833	2.48301

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-14.9158079346	2.0455416517	-7.292	3.34e-13 ***
AGE9	-0.8631792598	1.9179176926	-0.450	0.653678
AGE60	7.8953699347	1.1531062360	6.847	8.07e-12 ***
BRAND[T.ALLSPORT CHERRY SLAM]	-0.6764352636	0.0439775197	-15.381	< 2e-16 ***
BRAND[T.POWERADE TIDAL BURST]	-0.5668792833	0.0386292598	-14.675	< 2e-16 ***
EDUC	-0.4116364741	0.2545482023	-1.617	0.105889
ETHNIC	1.5640288388	0.3704432085	4.222	2.45e-05 ***

Feat	0.1809583378	0.0289361456	6.254	4.21e-10 ***
HHLarge	0.5758535352	1.1202837186	0.514	0.607248
HHSINGLE	-1.2994989915	0.6431985469	-2.020	0.043377 *
HVAL150	-0.0170106006	0.1160723589	-0.147	0.883489
INCOME	1.1081285530	0.1523858356	7.272	3.87e-13 ***
logprice	-1.9279757805	0.1022335815	-18.859	< 2e-16 ***
NOCAR	1.3682202340	0.3686223366	3.712	0.000207 ***
NWHITE	-2.9990492638	0.3601261975	-8.328	< 2e-16 ***
POVERTY	3.9248749278	1.4727328892	2.665	0.007713 **
REM	0.0109374237	0.0008148545	13.423	< 2e-16 ***
RETIRED	-4.2387109938	1.4713550006	-2.881	0.003977 **
SEASON[T.SPRING]	-0.5454425570	0.0326981122	-16.681	< 2e-16 ***
SEASON[T.SUMMER]	-0.1003607175	0.0326516708	-3.074	0.002121 **
SEASON[T.WINTER]	-0.3645311652	0.0329544873	-11.062	< 2e-16 ***
SINGLE	0.5251084984	0.8074069217	0.650	0.515475
STOREWEEK	0.0000041328	0.0000003286	12.579	< 2e-16 ***
UNEMP	13.9515194575	2.2781765170	6.124	9.54e-10 ***
WORKWOM	5.4854284624	1.1975824063	4.580	4.71e-06 ***
BRAND[T.ALLSPORT CHERRY SLAM]:Feat	0.1803015677	0.0548826450	3.285	0.001023 **
BRAND[T.POWERADE TIDAL BURST]:Feat	0.0958139509	0.0488175723	1.963	0.049715 *

Data Analysis

- How does prices vary across brands?

Rcmdr – Linear Model

Dependent variables [logprice]

Independent variables [BRAND] & [SEASON]

The screenshot shows the Rcmdr Linear Model dialog box on the left and its corresponding R console output on the right.

Linear Model Dialog Box:

- Enter name for model: LinearModel.1
- Variables (double-click to formula): Feat, flag, HHLARGE, HHSINGLE, HVAL150, INCOME
- Model Formula:
Operators (click to formula): + * : / %in% - ^ ()
Splines/Polynomials: (select variable and click) B-spline, natural spline, orthogonal polynomial, raw polynomial, deg. for polynomials: 5, df for splines: 5
- Subset expression: <all valid cases>
- Weights: <no variable selected>
- Buttons: Help, Reset, OK, Cancel, Apply

R Console Output:

```
Call:
lm(formula = logprice ~ BRAND + SEASON + STORE + Feat, data = Dataset)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.34368 -0.06384  0.00203  0.07061  0.27880 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.21616535  0.00329976 65.509 < 2e-16 ***
BRAND[T.ALLSPORT CHERRY SLAM] -0.04951889  0.00292890 -16.907 < 2e-16 ***
BRAND[T.POWERADE TIDAL BURST] -0.02377691  0.00274720 -8.655 < 2e-16 ***
SEASON[T.SPRING]           -0.00184857  0.00320456 -0.577  0.5641  
SEASON[T.SUMMER]            0.00526253  0.00237259  2.218  0.0266 *  
SEASON[T.WINTER]            0.02788266  0.00362552  7.691 1.63e-14 ***
STORE                  0.00015255  0.00002841  5.369 8.12e-08 *** 
Feat                   -0.17572544  0.00219462 -80.071 < 2e-16 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Data Analysis

- How does prices vary across brands?

ALLSPORT LEMON LIME: ($ACS=0, PTB=0$)

$\text{logprice} = 0.2162 - 0.0018 \times \text{SPRING} + 0.0053 \times \text{SUMMER} + 0.0279 \times \text{WINTER} + 0.0002 \times \text{STORE}$
– **0.1757***Feat

ALLSPORT CHERRY SLAM: ($ACS=1, PTB=0$)

$\text{logprice} = (0.2162 - 0.0495) - 0.0018 \times \text{SPRING} + 0.0053 \times \text{SUMMER} + 0.0279 \times \text{WINTER} + 0.0002 \times \text{STORE}$
– 0.1757*Feat
= **0.1667** - 0.0018*SPRING + 0.0053*SUMMER + 0.0279*WINTER + 0.0002*STORE
– **0.1757***Feat

POWERADE TIDAL BURST: ($ACS=0, PTB=1$)

$\text{logprice} = (0.2162 - 0.0238) - 0.0018 \times \text{SPRING} + 0.0053 \times \text{SUMMER} + 0.0279 \times \text{WINTER} + 0.0002 \times \text{STORE} -$
0.1757*Feat
= **0.1924** - 0.0018*SPRING + 0.0053*SUMMER + 0.0279*WINTER + 0.0002*STORE
– **0.1757***Feat

For all brands, price decreases if the store is more probable to offer on sale.

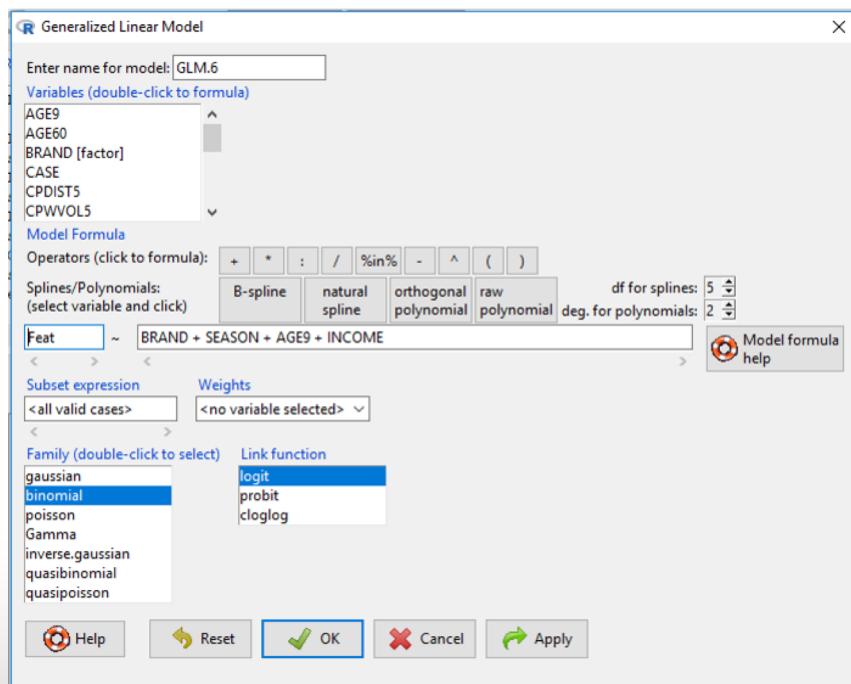
For the same number of store, the price of ALL is highest, of PTB is second highest, and of ACS is lowest.

In the same proportion of on sale, ACS has the lowest price while ALL have the highest price.

Data Analysis

- How does the proportion of times a brand is on sale vary across brand?

Rcmdr – Generalized Linear Model



```
Call:  
glm(formula = Feat ~ BRAND + SEASON + AGE9 + INCOME, family = binomial(logit),  
     data = Dataset)  
  
Deviance Residuals:  
    Min      1Q  Median      3Q      Max |  
-1.8590 -1.0304  0.6828  0.9789  1.7653  
  
Coefficients:  
              Estimate Std. Error z value Pr(>|z|)  
(Intercept)  2.69242   0.87585  3.074  0.00211 **  
BRAND[T.ALLSPORT CHERRY SLAM] 0.03568   0.07017  0.508  0.61112  
BRAND[T.POWERADE TIDAL BURST] -0.29990   0.06415 -4.675 0.00000294 ***  
SEASON[T.SPRING]   -0.85937   0.07111 -12.086 < 2e-16 ***  
SEASON[T.SUMMER]    0.86770   0.05650  15.357 < 2e-16 ***  
SEASON[T.WINTER]   -1.55389   0.08729 -17.801 < 2e-16 ***  
AGE9            -2.87401   1.03444 -2.778  0.00546 **  
INCOME          -0.17604   0.08332 -2.113  0.03462 *  
---  
Signif. codes:  0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.' 0.1 ' ' 1
```

Data Analysis

- How does the proportion of times a brand is on sale vary across brand?

ALLSPORT LEMON LIME: (ACS=0, PTB=0)

$$I = 2.692 - 0.859 \text{*SPRING} + 0.868 \text{*SUMMER} - 1.554 \text{*WINTER} - 2.874 \text{*AGE9} - 0.176 \text{*INCOME}$$

ALLSPORT CHERRY SLAM: (ACS=1, PTB=0)

$$I = 2.728 - 0.859 \text{*SPRING} + 0.868 \text{*SUMMER} - 1.554 \text{*WINTER} - 2.874 \text{*AGE9} - 0.176 \text{*INCOME}$$

POWERADE TIDAL BURST: (ACS=0, PTB=1)

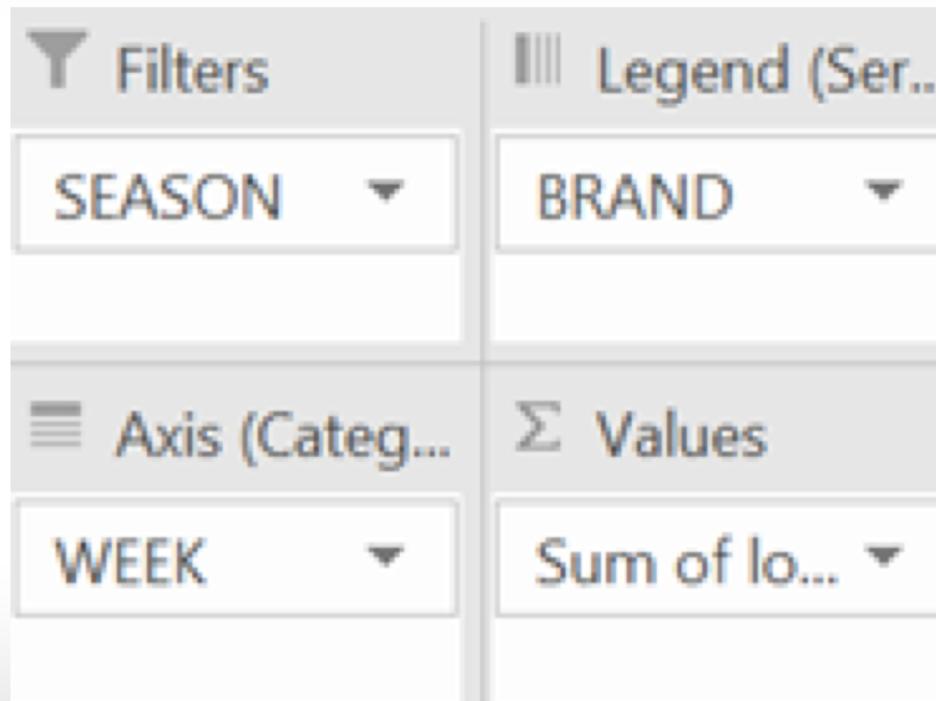
$$I = 2.392 - 0.859 \text{*SPRING} + 0.868 \text{*SUMMER} - 1.554 \text{*WINTER} - 2.874 \text{*AGE9} - 0.176 \text{*INCOME}$$

The probability of Feat = 1 (brand offered on sale) decreases if the store is in a higher income area. For the same level of income, ACS is most likely, ALL is second most likely, and PTB is least likely to be on sale.

Most factors are significant at least a 95% level.

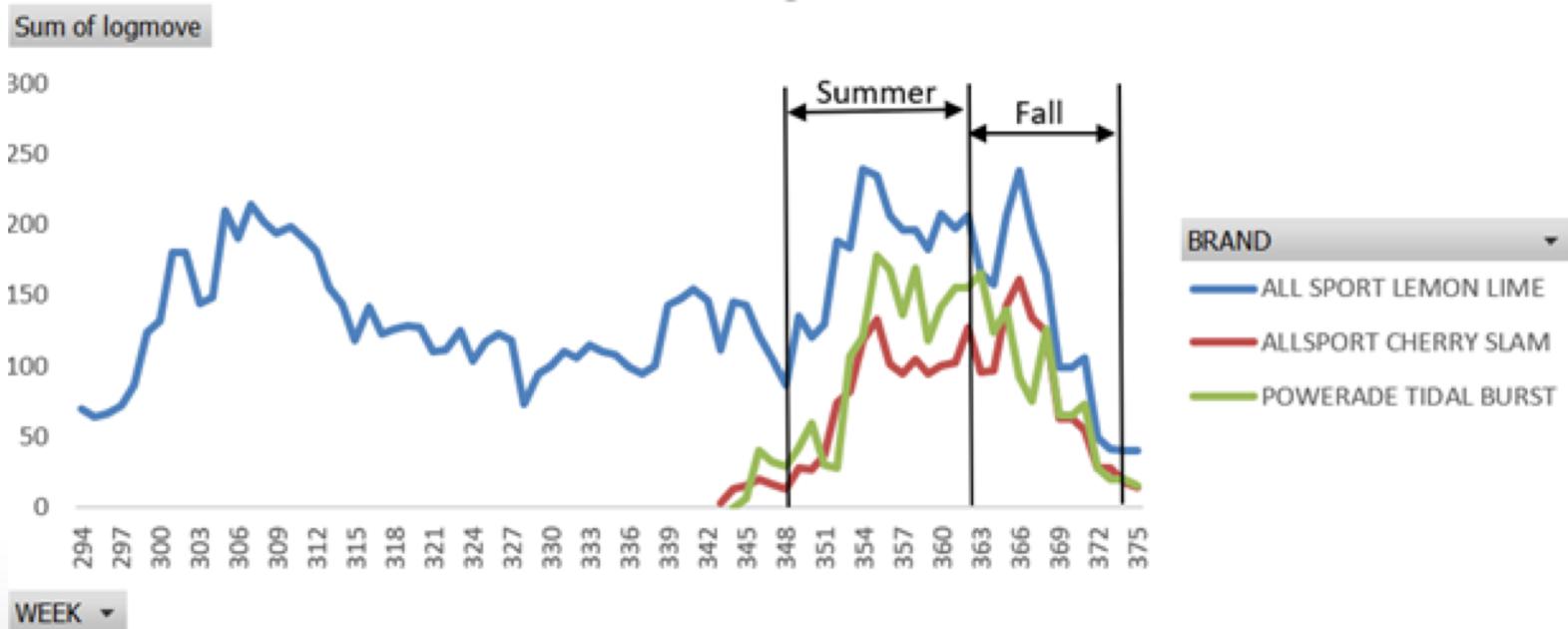
Data Analysis

- How does the demands for the three brands vary over time?



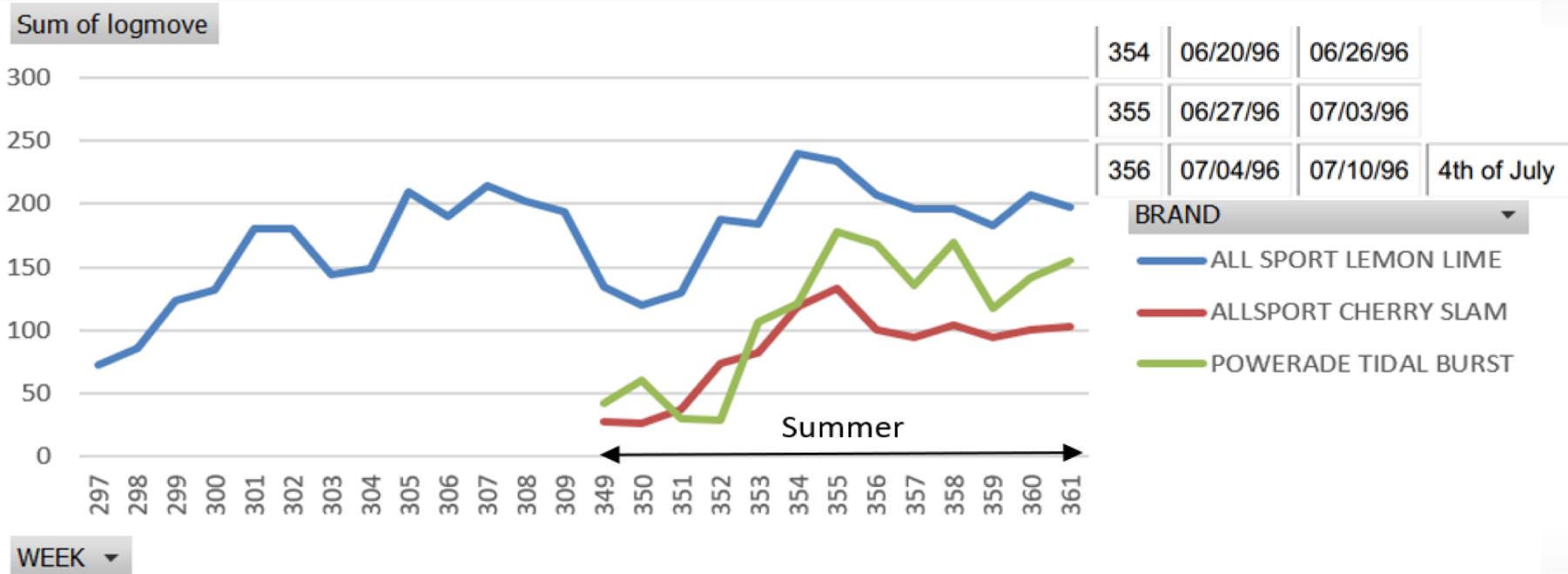
Data Analysis

- How does the demands for the three brands vary over time?



Data Analysis

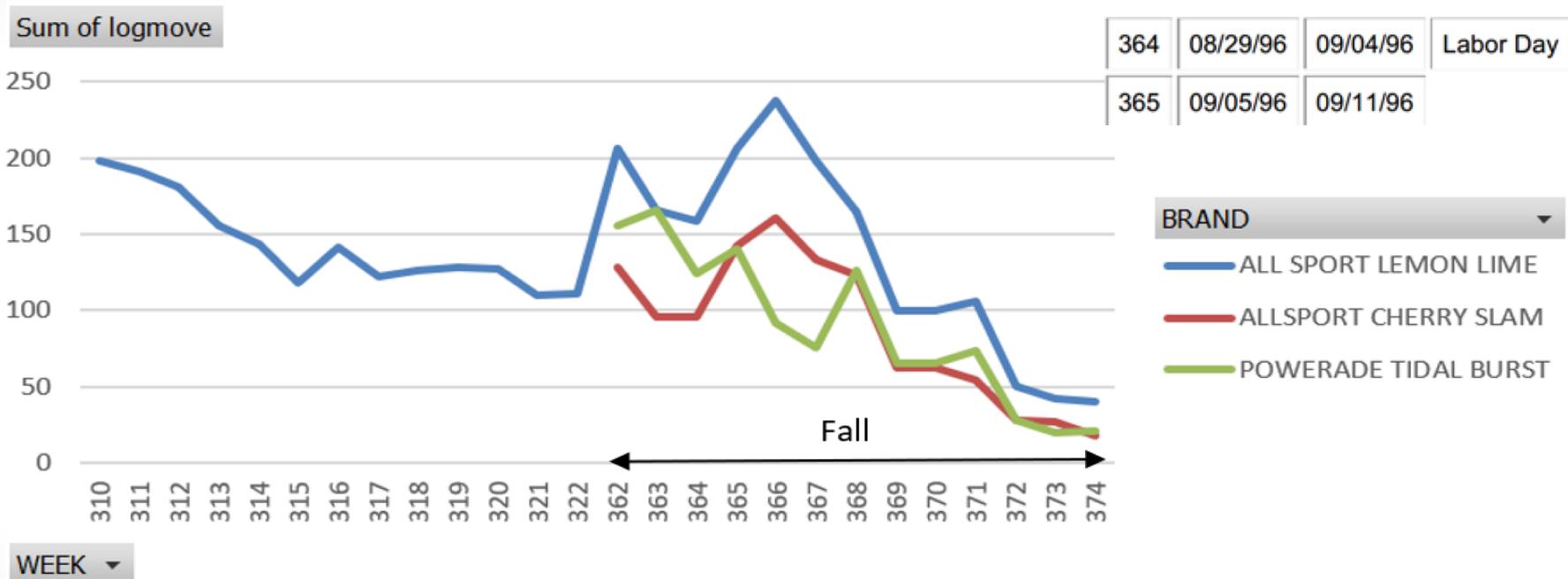
- How does the demands for the three brands vary over time?



The highest demand occurred during week 354 and 355
The Independence Day was in the week 356

Data Analysis

- How does the demands for the three brands vary over time?



Sales for All Sport Lemon Lime reached approximately 240 units
Sales for All Sport Cherry Slam approached 160 units

Data Analysis

- Develop a model, and test how it performs on a validation sample. For example, you can break your data randomly into two parts, estimation sample and validation sample, using Access (standard practice is to use two-thirds as the estimation sample, and other one-third as the validation sample). Estimate the model using the estimation sample, and assess how well the model predicts the dependent variable(s) in the validation sample.

Data Analysis

```
> randIndex <- sample(1:dim(SportsDrinks)[1])
> train_cutpoint2_3 <- floor((2*dim(SportsDrinks)[1])/3)
> testCutpoint <- dim(SportsDrinks)[1]-(train_cutpoint2_3+1)
> trainData <- SportsDrinks[randIndex[1:train_cutpoint2_3],]
> testData <- SportsDrinks[randIndex[train_cutpoint2_3+1:testCutpoint],]
> |  
  
> LMtrainData <- lm(logmove ~ AGE9 + AGE60 + BRAND + EDUC + ETHNIC + Feat + HHLARGE + HHSSINGLE + HVAL150 + INCOME + logprice + NOCAR + NWHITE + POVERTY + REM + RETIRED + SEASON + SINGLE +
+ STOREWEEK + UNEMP + WORKWOM + BRAND * logprice, data=trainData)  
  
> predict(LMtrainData,testData) -> a  
> View(a)
```



	row.names	x
1	556	2.3276207
2	7244	1.3021870
3	2415	2.1353120
4	7305	1.7525908
5	7818	1.7026336
6	2664	2.7393324
7	8129	1.5761021
8	8271	1.9397701
9	3953	1.2454479
10	7707	1.7888413
11	3053	1.6197723
12	2392	2.7861518
13	6030	2.0607570
14	4118	1.4794873
15	7421	2.0522302
16	5183	2.0776921
17	3829	2.2759121
18	1475	2.4530743
19	6637	2.3952870

Data Analysis

```
> compTable <- data.frame(testData[,11],a)  
> colnames(compTable) <- c('test','pred')  
> View(compTable)
```



	row.names	test	pred
1	556	2.3978953	2.3276207
2	7244	1.0986123	1.3021870
3	2415	1.0986123	2.1353120
4	7305	1.7917595	1.7525908
5	7818	1.3862944	1.7026336
6	2664	2.3025851	2.7393324
7	8129	2.7080502	1.5761021
8	8271	2.0794415	1.9397701
9	3953	0.0000000	1.2454479
10	7707	0.0000000	1.7888413
11	3053	0.0000000	1.6197723
12	2392	3.0910425	2.7861518
13	6030	2.1972246	2.0607570
14	4118	0.0000000	1.4794873
15	7421	2.9957323	2.0522302
16	5183	2.1972246	2.0776921
17	3829	2.6390573	2.2759121
18	1475	2.6390573	2.4530743
19	6637	3.5553481	2.3952870

Data Analysis

```
> compTable$error <- compTable$test - compTable$pred  
> View(compTable)
```



	row.names	test	pred	error
1	556	2.3978953	2.3276207	0.07027459570
2	7244	1.0986123	1.3021870	-0.20357472011
3	2415	1.0986123	2.1353120	-1.03669971551
4	7305	1.7917595	1.7525908	0.03916871697
5	7818	1.3862944	1.7026336	-0.31633920779
6	2664	2.3025851	2.7393324	-0.43674730829
7	8129	2.7080502	1.5761021	1.13194808347
8	8271	2.0794415	1.9397701	0.13967141380
9	3953	0.0000000	1.2454479	-1.24544793559
10	7707	0.0000000	1.7888413	-1.78884125642
11	3053	0.0000000	1.6197723	-1.61977234277
12	2392	3.0910425	2.7861518	0.30489065370
13	6030	2.1972246	2.0607570	0.13646760756
14	4118	0.0000000	1.4794873	-1.47948728324
15	7421	2.9957323	2.0522302	0.94350202586
16	5183	2.1972246	2.0776921	0.11953247620
17	3829	2.6390573	2.2759121	0.36314523764
18	1475	2.6390573	2.4530743	0.18598300930
19	6637	3.5553481	2.3952870	1.16006105256

