Sales Difference of Orthopedic Equipments Among US States

He, Chunsheng
Liu, Lu
Lou, Jianxiong
Wang, Jue
Zhang, Hong
Objective

- The objective of this study is to fit a model to the sales data
- The model should demonstrate the differences between states
Data Description

- **STATE**: STATE NAME
- **BEDS**: NUMBER OF HOSPITAL BEDS
- **RBEDS**: NUMBER OF REHAB BEDS
- **OUT-V**: NUMBER OF OUTPATIENT VISITS
- **ADM**: ADMINISTRATIVE COST (In $1000's per year)
- **SIR**: REVENUE FROM INPATIENT
- **SALEY**: SALES OF REHABILITATION EQUIPMENT SINCE JAN 1
- **SALES12**: SALES OF REHAB. EQUIP. FOR THE LAST 12 MO
- **HIP95**: NUMBER OF HIP OPERATIONS FOR 1995
- **KNEE95**: NUMBER OF KNEE OPERATIONS FOR 1995
- **TH**: TEACHING HOSPITAL? 0, 1
- **TRAUMA**: DO THEY HAVE A TRAUMA UNIT? 0, 1
- **REHAB**: DO THEY HAVE A REHAB UNIT? 0, 1
- **HIP96**: NUMBER HIP OPERATIONS FOR 1996
- **KNEE96**: NUMBER KNEE OPERATIONS FOR 1996
- **FEMUR96**: NUMBER FEMUR OPERATIONS FOR 1996
Step 1--New Data Set

STATE : State Name
SALES12 : Average Sales Of Rehab. Equip. For The Last 12 Months
COUNT : Number of Hospitals in a State
BEDS : Average Number Of Hospital Beds
RBEDS : Average Number Of Rehab Beds
OUT-V : Number Of Outpatient Visits Per Hospital
ADM : Administrative COST(In $1000's Per Year) Per Hospital
SIR : Revenue From Inpatient Per Hospital
HIP95 : Average Number Of Hip Operations For 1995
KNEE95 : Average Number Of Knee Operations For 1995
HIP96 : Average Number Hip Operations For 1996
KNEE96 : Average Number Knee Operations For 1996
FEMUR96 : Average Number Femur Operations For 1996
TH : Average Number of Teaching Hospital
TRAUMA : Average number of Trauma Unit
REHAB : Average Number of Rehab Unit
Step 2—Data Transformation

**count**

```
 0 10
0 200 400
```

**beds**

```
0 10
100 300
```

**rbeds**

```
0 10
0 4 8 12
```

**out**

```
0 10 20
0 50000 150000
```

**adm**

```
0 10
0 2000 8000 14000
```

**sir**

```
0 10
2000 6000 10000
```

**salesy**

```
0 10 20
0 5000 15000
```

**sales12**

```
0 10 20
0 40 80 120
```

**hip95**

```
0 5 10
10 30 50 70
```

**knee95**

```
0 5 10
0 4 8 14
```

**th**

```
0 10
0.0 0.4 0.8
```

**trauma**

```
0 5 10
0.0 0.2 0.4
```

**rehab**

```
0 5 15
0.0 0.2 0.4
```

**hip96**

```
0 5 15
20 40 60 80
```

**knee96**

```
0 5 15
0 4 8 14
```

**femur96**

```
0 5 15
20 40 60 80
```
Step 2—Data Transformation
(cont’d)
Step 3—Data Grouping

- We separate the data set into two groups according to
  - State average sales
  - Clustering method (Kmeans)
Step 3—Data Grouping (cont’d)

- Method of state average sales
  - 24 states above average sales (group 1)
  - 28 states below average sales (group 0)

- Kmeans method
  - 22 states in group 1
  - 30 states in group 0
  - Difference: SC and WI
Step 4—Logistic Regression

- Dependent variable
  - 1—group 1
  - 0—group 0
  - Group ~ Bernoulli(p)

- Independent variables
  - Total variables available: 14
  - Model used (stepwise): 5
Step 4—Logistic Regression (cont’d)

```
summary(glm(formula = group ~ beds + out + trauma + hip96 + knee96, family = "binomial"))
```

**Summary**

<table>
<thead>
<tr>
<th>Deviance Residuals:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
</tr>
<tr>
<td>-1.7464</td>
</tr>
<tr>
<td>0.7694</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>beds</td>
</tr>
<tr>
<td>out</td>
</tr>
<tr>
<td>trauma</td>
</tr>
<tr>
<td>hip96</td>
</tr>
<tr>
<td>knee96</td>
</tr>
</tbody>
</table>

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 71.779 on 51 degrees of freedom
Residual deviance: 51.372 on 46 degrees of freedom
AIC: 63.372
Step 4—Logistic Regression (cont’d)

- Residual vs. Fitted
Step 4—Logistic Regression (cont’d)

- Residual against time
Step 5—Data Regrouping

- The difference in grouping between data and model

<table>
<thead>
<tr>
<th>State</th>
<th>Sales12</th>
<th>Current group</th>
<th>Predicted Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>IN</td>
<td>31.853448</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MN</td>
<td>18.273684</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MO</td>
<td>33.84252</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NE</td>
<td>28.537037</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RI</td>
<td>4.692308</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>103.8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GA</td>
<td>45.065574</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MT</td>
<td>49.62963</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NV</td>
<td>78.894737</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NY</td>
<td>43.966805</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>SC</td>
<td>39.508197</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>WI</td>
<td>38.393162</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Step 6—Alternative model

- Prediction accuracy of previous model is only about 77%
- Use Kmeans clustering
  - 22 states in group 1
  - 30 states in group 0
Step 6—Alternative model (cont’d)

```
glm(formula = kgroup ~ beds + rbeds + out + sir + knee95 + trauma +
    hip96 + femur96, family = "binomial", data = sm)

Summary
Deviance Residuals:
  Min        1Q    Median        3Q       Max
-2.64007  -0.60411  -0.05059   0.43079   1.90822

Coefficients:
  Estimate Std. Error z value Pr(>|z|)
(Intercept)  257.726    110.443   2.334  0.01962 *
beds          12.384      5.967   2.075  0.03794 *
rbeds         -8.180      4.316  -1.895  0.05803 .
out         -159.469     65.212  -2.445  0.01447 *
sir          -67.589     39.427  -1.714  0.08648 .
knee95       109.082     41.990   2.598  0.00938 **
trauma        20.284      7.224   2.808  0.00499 **
hip96         -5.370      2.432  -2.208  0.02722 *
femur96       15.192      8.363   1.817  0.06928 .

Null deviance: 70.852  on 51  degrees of freedom
Residual deviance: 37.560  on 43  degrees of freedom
AIC: 55.56
```
Step 6—Alternative model (cont’d)
Step 6—Alternative model (cont’d)

- The difference in grouping between data and alternative model

<table>
<thead>
<tr>
<th>State</th>
<th>Sales12</th>
<th>Current group</th>
<th>Predicted group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>27.975</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IN</td>
<td>31.85345</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MN</td>
<td>18.27368</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MO</td>
<td>33.84252</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NE</td>
<td>28.53704</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RI</td>
<td>4.692308</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>103.8</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Conclusions

The company’s sales in the states of IN MN MO NE and RI are currently lower than the national average. Based on the company’s sales record across the US, our statistical analysis has concluded that the company’s sales in these five states can be increased to national average level. There are 405 hospitals in these five states, should appropriate measure be taken, the company’s sales will be increased by 4121.
Conclusions (cont’d)

States of interests, Original model

State of interests, Model 2
Conclusions (cont’d)