VANDERBILT UNIVERSITY

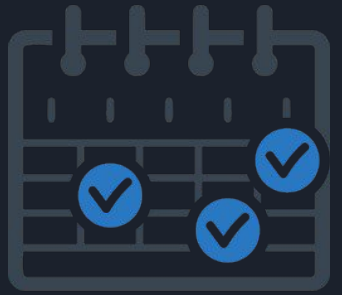
MEDICAL CENTER

O.R. Predictive Analytics

By: Amrita Nair and Jesse Pacheco

Process to choose target scheduling date

- Find the benchmark - Our benchmark is the mean percentage increase in surgeries between target scheduling day (T-28, T-21, T-14) and date of surgery.
- Develop multiple Regression models for target scheduling days
  - ➢ Input: DOW, SurgDate, target scheduling day
  - ➢ Output: Predicted Number of Surgeries
- Compare Model Performance against the benchmark
- Choose target scheduling date with the best cost-benefit



We chose T-14 as the target scheduling date

	Linear Regression Model				Benchmark	
Target	T-stat	Adj R2	MAPE	RMSE	MAPE	RMSE
T - 28	6	21%	7.6%	10.7	20%	30
T - 21	8	30%	7.4%	10	15%	21
T - 14	10	43%	6.2%	9	11%	16
T - 12	11	48%	6%	8.7	11%	14.6
T - 10	12	51%	5.5%	8.2	9%	13
T - 8	15	60%	4.6%	7.3	7%	9

We chose 2 weeks prior actual surgery date as the target scheduling date

- If target schedule date is 14 days ahead:
  - ➢ On average our model is off by 6.2% i.e. by 9 surgeries
  - ➢ Better than using benchmark model by 7 surgeries
  - ➢ 43% of the variance is explained and this date is statistically significant.
- If target schedule date is between 10 to 14 days prior, on average our model is off by more than 8 surgeries.
- If target schedule date is between 14 to 28 days prior, on average our model is off by more than 10 surgeries.

Improving Resource Allocation and Efficiency using our model

- **Staff** - Our model accounting for daily variations would decrease daily residuals. VUMC can allocate staff (in case of limited resources) on weekends
- **Shifts** - Utilizing our model, staffing can be scheduled 14 days in advance of surgery date with minor modifications being made 24 hours before day of surgery.
- **Labor Expense** - Decrease in overtime pay for nurses and increased research productivity from anesthesia department. Nursing staff can decide when to apply holiday leave.
- **Facilities and Equipment** - More accurate gauge for closing OR and placing sterile equipment orders with third party supplier.

Feedback/Questions

APPENDIX

Process to predict using benchmark

- Percentage increase between target schedule date ‘T-14’ and Actuals
- Calculated benchmark = mean of percentage increase
- Predict using this benchmark.
- Find error between Actuals and Predictions - MAPE and RMSE

Our Code to Predict using Benchmark

```r
#Data Preperation---------------------------------
library(readr)
vumc = read_csv("C:/Users/amrit/Downloads/VUMC Case Data.csv")
vumc = vumc[vumc$Actual > 30,]
vumc$SurgDate = as.Date(vumc$SurgDate, "%m/%d/%Y")
vumc$DOW = factor(vumc$DOW, levels = c("Mon", "Tue", "Wed", "Thu", "Fri"))

#Partitioning data--------------------------------
N = nrow(vumc)                        #No of rows
trainingCases = sample(N,round(N*0.6))  #indices of training cases
train = vumc[ trainingCases,]         #60% of data used for training
test  = vumc[-trainingCases,]         #40% of data used for testing


#Benchmark prediction-----------------------------------
benchmark_t14 = (train$Actual - train$`T - 14`)/train$`T - 14` #%age increase
avg_t14 = mean(benchmark_t14)                                  #Mean %age increase
predicted_bench_t14 = ceiling((test$`T - 14`*avg_t14)+test$`T - 14`) #Predict using benchmark
errors = test$Actual - predicted_bench_t14                     #Find errors in predicted benchmark
hist(errors,col = "blue",xlab="Error in predicted benchmark 14 days ahead")
mape = mean(abs(errors/test$Actual))    #On average this model is 10.21% off
rmse = sqrt(mean(errors^2))             #On average this model is 16 surgeries off
```

Choosing our ML model

- Actual = bo + b1 x ‘T-14’
  On average this model was off by 5.9% or 8.8 surgeries
  Adj R2 explained 41.5% of variations and ‘T-14’ is very significant

- Actual = bo + b1 x ‘T-14’ + b2 x DOW
  On average this model was off by 6.2% or 8.9 surgeries
  Adj R2 explained 42% of variations and DOW is very insignificant

- Actual = bo + b1 x ‘T-14’ + b2 x ‘T-10’
  On average this model was off by 5.6% or 8.9 surgeries
  Adj R2 explained 42% of variations. Both are significant

- Actual = bo + b1 x ‘T-10’
  On average this model was off by 5.7% or 8.37 surgeries
  Adj R2 explained 51.5% of variations.

Our Code to Predict using Linear Regression

```
#Machine Learning Prediction-------------------------------
model = lm(Actual ~ `T - 14`, data = train)
predictions = predict(model,test)
observations = test$Actual
errors = observations - predictions
hist(errors)
mape = mean(abs(errors/observations))   #On average, this model is 5.9% off
rmse = sqrt(mean(errors^2))             #On Average, this model is 8.8 surgeries off
summary(model)                          #Adj R2 = 41.5%; T-14 very significant

model = lm(Actual ~ DOW + `T - 14`, data = train)
predictions = predict(model,test)
observations = test$Actual
errors = observations - predictions
hist(errors)
mape = mean(abs(errors/observations))   #On average, this model is 6.2% off
rmse = sqrt(mean(errors^2))             #On Average, this model is 8.9 surgeries off
summary(model)                          #Adj R2 = 42%; DOW is insignificant


model = lm(Actual ~ `T - 14` + `T - 10`, data = train)
predictions = predict(model,test)
observations = test$Actual
errors = observations - predictions
hist(errors)
mape = mean(abs(errors/observations))   #On average, this model is 5.6% off
rmse = sqrt(mean(errors^2))             #On Average, this model is 8.3 surgeries off
summary(model)                          #Adj R2 = 51.8%;
```