UCL
SCHOOL OF
MANAGEMENT

UCL

MSIN0095: Operations Analytics

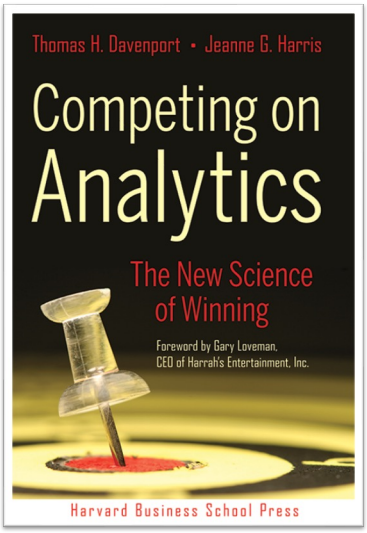
Class 1-4: Process Analysis
Class 5,7: Waiting Time Analysis
Class 6: Inventory Management – Newsvendor Model
Class 8: Inventory Management – Newsvendor, Periodic Review
Class 9: Inventory Management – EOQ
Class 10: Inventory Management – Amazon Distribution Strategy
Class 11: Supply Chain Management I: Beer Game
Class 12: Supply Chain Management II
Class 13: Supply Chain Management III: Strategic Sourcing, Sustainable Supply Chains
Class 14: Demand Forecasting I
Class 15: Demand Forecasting II – Caesars Entertainment

Class 16-17: Revenue Management

1

Competing on Analytics

- Some high-performing firms are building their competitive strategies around data-driven insights.
- Examples of analytics competitors are:
  - Capital One, Barclays (Finance)
  - Boston Red Sox, Oakland A's (Sports)
  - Amazon (Retail)
  - Proctor & Gamble (CPG)
  - Caesars Entertainment (Gaming)
  - Marriott International (Hotel)



2

2

3

3

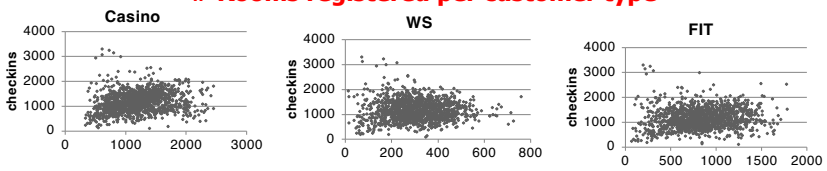
Multiple regression model

4

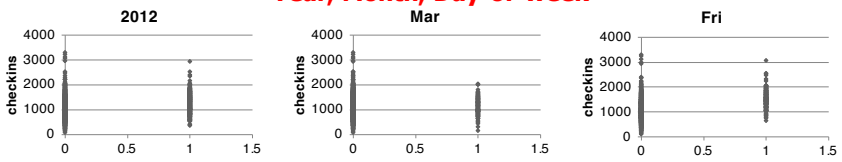
4

Preliminary analysis

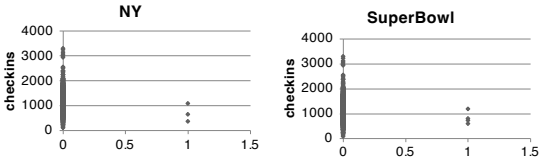
 # Rooms registered per customer type



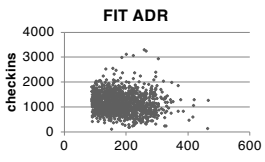
Year, Month, Day-of-week



Holidays and Events



Room rate



5

5

Correlation Matrix and Stepwise Regression

	MLK	Pres Day	Sun	Tue	Wed	Fri	2010	2012	Casino	SE	Seg 0-3	FIT	Group	FIT ADR
checkins	0.11	0.18	0.24	-0.20	-0.15	0.36	-0.11	0.20	0.21	0.17	0.16	0.12	-0.12	-0.11

Structural multicollinearity

	CNY7Days	Feb
CNY15Days	0.680508339	0.444570474

Data-based multicollinearity

	SE (Special event)	Seg 0-3 (Cust. Seg.)	WS (Wholesale)	FIT ADR (Room rate)
Casino	0.793460189	0.818783647		
SE		0.775171344		
FIT			0.493576989	
Group				0.684225847

Why is multicollinearity bad?

- Coefficients β_k depends on which variables are included
- Precision of coefficients decrease with more variables
- Hypothesis tests for β_k=0 may yield different conclusions

6

Forward/Backward Selection

- Forward Selection:
  - Start by choosing the independent variable that explains the most variation in the dependent variable
  - Add new independent variable that explains the most residual variation
  - Repeat until no variables "significantly" explain residual variation
- Backward Selection:
  - Start with all the variables in the model, and drop the least "significant", one at a time until you are left with only "significant" variables.
- Mixture of the two:
  - Perform a forward selection but drop variables that become no longer "significant" after introducing new variables.

Regression Statistics

Multiple R	0.623236
R Square	0.388423
Adjusted R Square	0.38085
Standard Error	321.5244
Observations	1309

ANOVA

	df	SS	MS	F	Significance F
Regression	16	84829124	5301820	51.2858	1.5E-125
Residual	1292	1.34E+08	103377.9		
Total	1308	2.18E+08			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	152.8	150.9	1.0	31.159%	-143.3	448.9
Sun	488.4	29.3	16.7	0.000%	430.8	545.9
Mon	230.4	30.3	7.6	0.000%	170.9	289.8
Wed	86.6	29.8	2.9	0.375%	28.1	145.1
Thu	204.7	29.2	7.0	0.000%	147.4	261.9
Fri	566.3	29.7	19.1	0.000%	508.1	624.6
NY	-598.7	187.4	-3.2	0.143%	-966.3	-231.0
MLK	839.0	162.7	5.2	0.000%	519.9	1158.1
Pres Day	1327.7	163.9	8.1	0.000%	1006.1	1649.2
15th of Month	164.5	56.2	2.9	0.350%	54.2	274.9
Casino	0.2	0.0	4.4	0.001%	0.1	0.3
FIT	0.4	0.1	5.9	0.000%	0.2	0.5
Group	0.2	0.0	5.0	0.000%	0.1	0.3
FIT ADR	-0.6	0.2	-2.8	0.484%	-1.0	-0.2
2010	110.6	38.2	2.9	0.382%	35.7	185.4
2011	156.6	38.1	4.1	0.004%	81.9	231.4
2012	214.4	28.8	7.5	0.000%	157.9	270.8

Forecast accuracy

	Mean Percentage Error	Mean Absolute Error	Mean Absolute Percentage Error
Moving average	-2%	225	25%
Decomposition	-3%	184	19%
Multiple regression	-14%	280	34%

9

9

Pros and Cons of models

	Pros	Cons
Moving Average	• Simple and intuitive, only need to keep recent data	• Not responsive to earlier trends
Time Series Decomposition	• Still simple (although not as simple as moving averages), • Captures trend, cycle, seasonal effect	• Needs to re-estimate trend line, cycle, seasonal effect every so often
Multivariate Regression	• Clear managerial implications • Identifies key drivers of demand (seasonality, which holiday, customer type, price) • Quantifies their impacts separately	• Complicated to maintain, needs lots of data that might not be available 2 weeks in advance • Also needs to re-estimate every so often to capture changes in coefficients (rolling regression)

10

10

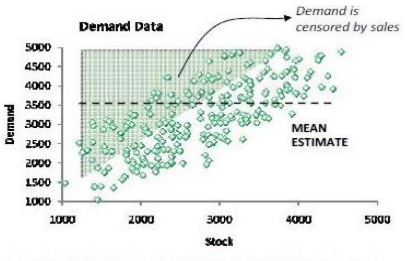
Caesars Entertainment: Takeaways

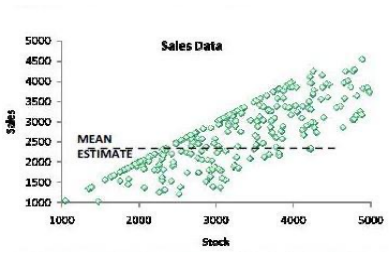
- Analytics competitors use data to improve business functions through using sophisticated quantitative techniques.
- How can we use data to predict demand?
  - Time series methods: Find historical patterns to make predictions
  - Regression methods: Model the cause-and-effect relationships of demand
- Out-of-sample validation (withholding data) is a good way to realistically compare forecasting performance of different models.

Another way to forecast?

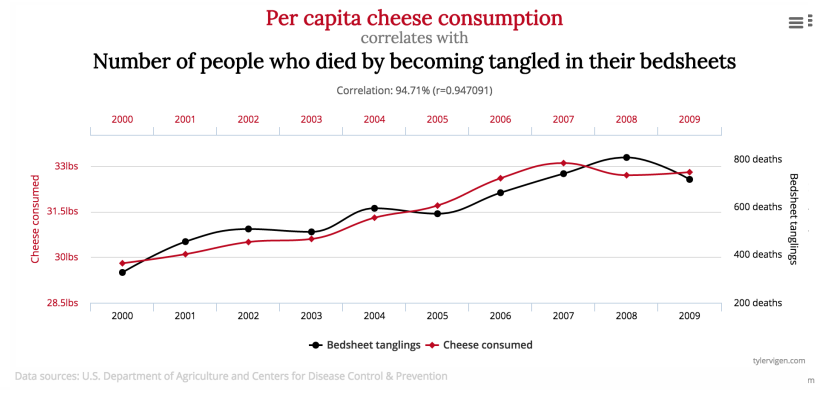
Method 1: Analytics Methods

- Pros/Cons?
  - Assumes structural stability, ignores “inflection points” in customer behavior
  - Does not work for new products, or where historical data does not exist
  - Often demand is not observed, only sales





Run Regression on What Again?!



Method 2: The Quaker Method

- Pros/Cons
  - Works for new products, incorporates human judgement
  - Organizational issues: politics around the table, anonymity
  - Difficult to account for level of confidence/information of experts.

Method 3: The Quaker Method on Steroids

- Poll a large number of people, the crowd! (or create prediction market!)
- Pros/Cons
- Works well for new feature development, avoids political issues.
- Works if someone's decisions are independent of everyone else'
- Potential for herding effect.
- Prediction markets are no crystal balls, only aggregate information.