Customer Segmentation

Doug Gray

DataScience@SMU

Customer Segmentation and Dynamic Pricing

- Airlines
- Movie theaters
- Sports teams
- Barber shops
- Credit card/retail banking companies

Catalina Marketing Targeted Offers

Offers are ***10 times*** more likely to be acted upon when backed up by analytics.

DataScience@SMU

Harrah's Downward Death Spiral

Doug Gray

Harrah’s Downward Death Spiral

- Losing customers
- Decreasing market share
- Decreasing revenues and profits
  - Failed to meet targets from 1991–97
- Aging properties
- Next stop: chapter 11 or 7
- Analytics to the rescue!

Harrah's Executive Leadership Profile

- Gary Loveman, PhD (Economics), MIT
  - Harvard Business School (HBS) professor
  - Consultant and speaker
  - Harrah's consultant, COO, then CEO
  - An academic theoretician who made the transition to corporate leadership
- ***Paper: "Putting the Service-Profit Chain to Work"***
  - The paper focused on the relationship between company profits and customer loyalty, and the importance of rewarding employees who interact with customers

Harrah’s Near Death Experience

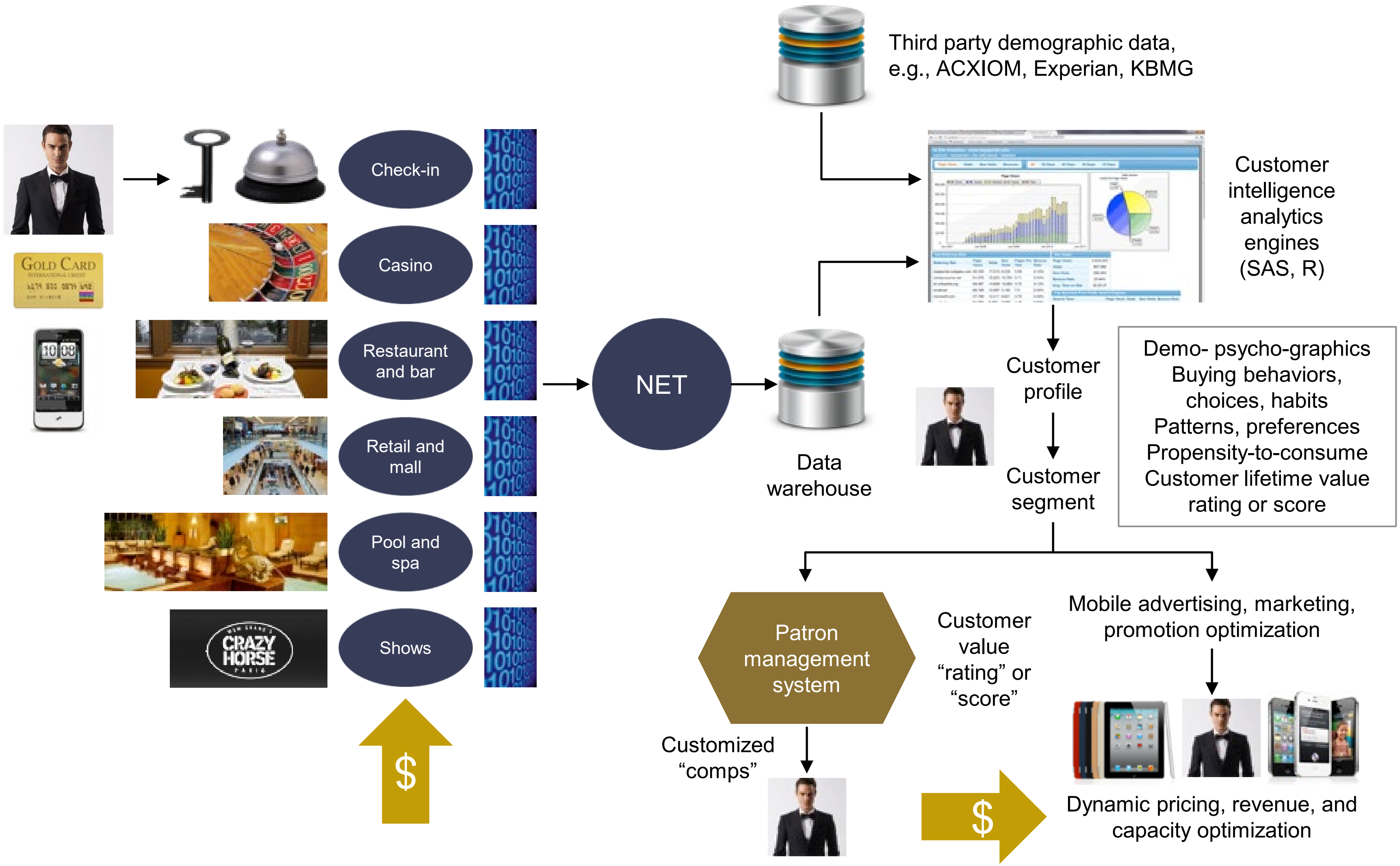
Market share growth

36% (1998) → 43% (2004)

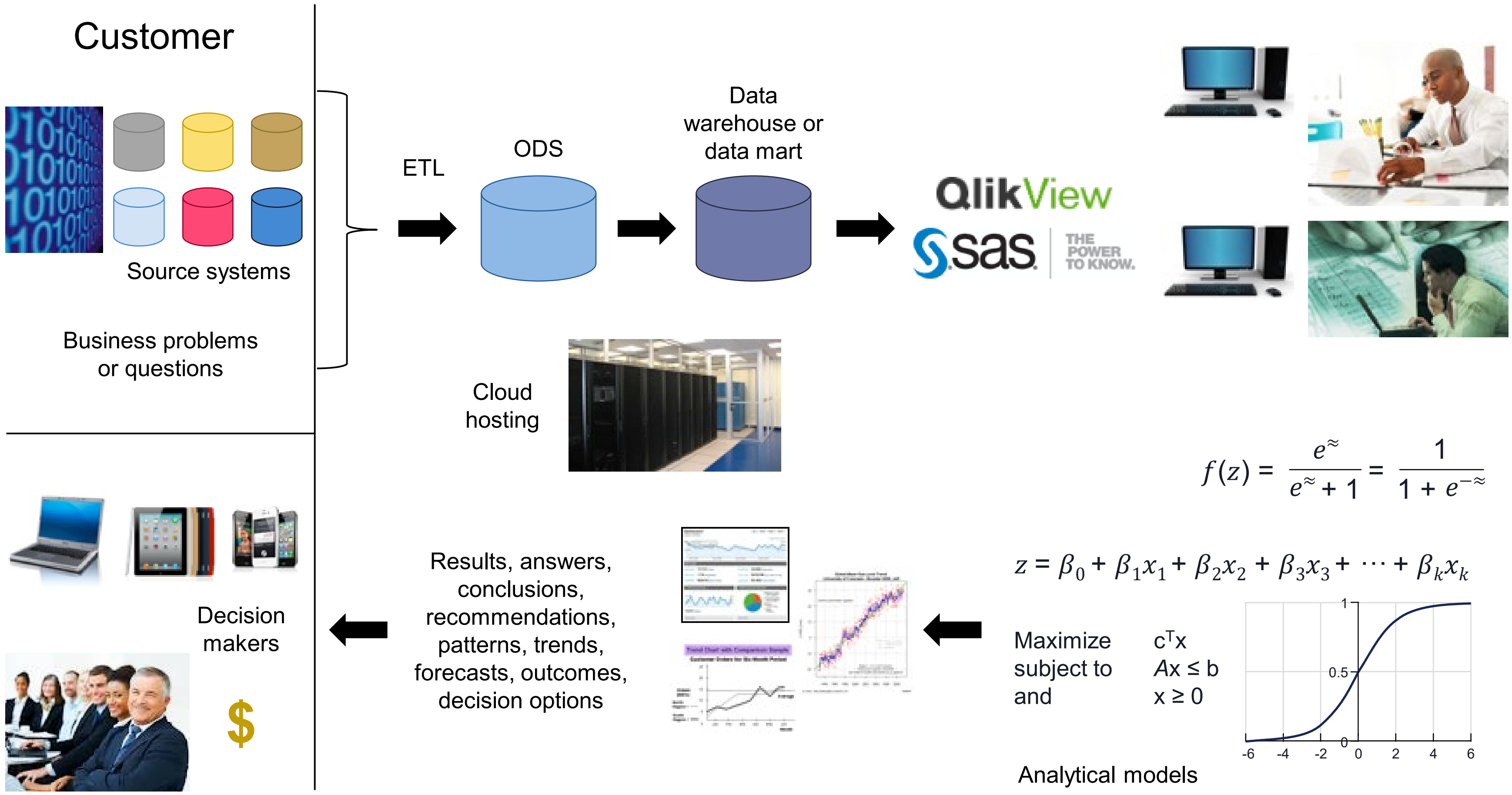
Same store sales gains 23/24 quarters

Failed to meet revenue and profit (1991-97)

Analytics Architecture



Analytics Architecture Flow



DataScience@SMU

Analytics Executive Profile

Gary Loveman

Doug Gray

Analytics Executive Profile: Gary Loveman

- "I am purely empirical."
  - *"I see my customers as a set of probabilities wrapped in human flesh."* ;)
  - *"I am not attached to any romantic notion of how this business should be run. I am only driven where the evidence takes me."* ***(Let the data speak!)***

Analytics Executive Profile: Gary Loveman

- Gary's method of customer profiling: segmentation
  - Gender and age: female and over 50 (most lucrative demographic)
    - Lots of "small fry"
    - Not a few "whales"
  - Games played and amount wagered
  - Demographics
    - Where we live
    - What we do (leisure activities and how we make a living)
    - Income level

Analytics Executive Profile: Gary Loveman

- *“Mathematics is the language of logical expression and my mind works very logically. I have found in the casino business that is a great help most of the time.”*
  - No HiPPOs allowed!
  - No managing by gut instinct
- Built Harrah’s from 15 regional casinos to 39 in the U.S. and 13 overseas
- *“My sort of logic excels at things like acquisitions or expansions or developments where I can look at the numbers and make it a defined, deductive problem,”*
  - Acquired World Series of Poker, Binion’s Horseshoe, Caesars, Imperial Palace, Planet Hollywood Resort and Casino, and golf course in Macau (world’s most lucrative market)

Analytics Executive Profile: Gary Loveman

- Customer targeting
  - Whales vs. grandma and grandpa
  - Customer database
  - Targeted offers based on demographic and gaming transactional history
  - Come back soon inducements
- Gaming revenue accounted for 80% of Harrah’s $9 billion revenue vs. 45% industry
- Food, beverage, retail, conferences, and trade shows now account for more revenue than gaming since the economic crash in 2008–09

Analytics Executive Profile: Gary Loveman

- Sometimes the numbers alone fail Loveman in his decision-making.
    - He passed on buying Macau gaming license for $900 million (one of only six available).
    - *“There are two scarce resources in Macau; land and gaming licenses.”*
    - *“The quantification took me in the opposite direction. You had to have a kind of intuitive courage, and I am not well-suited to those kinds of decisions.”*
    - *“Big mistake. I was wrong, I was really wrong.”*

Analytics Executive Profile: Gary Loveman

- Macau is now the world's largest gaming market.
  - Las Vegas Sands and Wynn Resorts are now making more money there than in Vegas.
  - Loveman spent $600 million on a Macau golf course, in hopes of getting a gaming license later.
- *"I try to be self-reflective of the fact that this is a short-coming of mine. I am trying to address that vulnerability by asking myself over and over again if I am over-thinking the problem. Am I missing something?"*

Analytics Executive Profile: Gary Loveman

- Harrah's **Total Rewards** loyalty program enabled annual double-digit growth for five years in a row 2003–7.
  - Far more profitable to retain loyal, repeat customers rather than have to market to and secure new customers
- ***Paper: "Putting the Service-Profit Chain to Work" (1994)***
  - The paper focused on the relationship between company profits and customer loyalty and the importance of rewarding employees who interact with customers
  - Loveman pioneered the use of analytics in casinos and quantified the link between happier customers and more profitable companies

Analytics Executive Profile: Gary Loveman

- Casinos are largely a *commodity business* with little opportunity for differentiation between slot machines and gaming tables
  - *Customer service and targeted marketing* can make a significant difference in company performance
- **Total Rewards**
  - _____ million members by 2010 with a $_____ million annual IT budget
  - Described as the linchpin of Harrah's success
  - The first loyalty program to be applied across every casino in a company
  - Allowed for more accurate analysis of betting patterns
  - More equitable distribution of "comps," i.e., free rooms, meals, show tickets

Analytics Executive Profile: Gary Loveman

- **Total Rewards** led Loveman to a startling conclusion—*Harrah's makes more money from elderly slot machine players than any other demographic, even high-roller multi-millionaire "whales."*
  - *"The slot player was the forgotten customer and the slot player offered us the biggest benefit in terms of data collection. I had to be willing to be unsexy in this. I can take you to a casino that would have a lot of young beautiful people in there and you would say, 'Man, this is a happening place.' I could take you to another place where there are a lot of people who look like your parents. The latter would be more profitable than the former. My job is to make the latter."*
  - *"The motor" of the Harrah's juggernaut "is the database."*

Analytics Executive Profile: Gary Loveman

- Proprietary software assigns a ***value*** to each Total Rewards member
  - The ***expected amount of money you will lose per visit to a Harrah's casino***, based on the odds of the games you play, how long you play, and how much you wager
  - 30-year old female slot player who has made 36 trips to Harrah's Chicagoland casino, playing an average of 47 minutes losing $156 per visit; an active buffet customer, she will be enticed back with buffet and slot machine offers

DataScience@SMU

The $529 Sucker

Doug Gray

DataScience@SMU

The $529 Sucker

- Gambled for exactly 8 hours, 34 minutes
- For example, sat down to play Pai Gow poker at 12:39 PM, got up one hour and six minutes later; bought in for $800 and cashed out for $1,300
- Tracked where he had eaten, how much he had spent, where he shopped
- Based on his age, gender, zip code, and gaming preferences, a value of $529 was assigned, i.e., *how much he was expected to lose every time he visited a Harrah's property*
- ***"You will be getting an attractive offer from us. Please keep playing."***

Gambling Industry and Ethics

- According to the American Gaming Association, gambling is legal in 40 U.S. states and is a $240 billion industry employing 1.7 million people contributing $8.5 billion in state and local tax revenues.
- Critics claim gambling leads to increased political corruption, compulsive gambling, and increased crime rates.
- Gambling is viewed as another form of entertainment, no different than theater, concert and sporting events tickets, or amusement parks.

DataScience@SMU

Harrah's Cherokee Casino

Revenue Management

Doug Gray

Harrah’s Cherokee Casino and Hotel

A “killer app” for revenue management

Revenue Management Overview

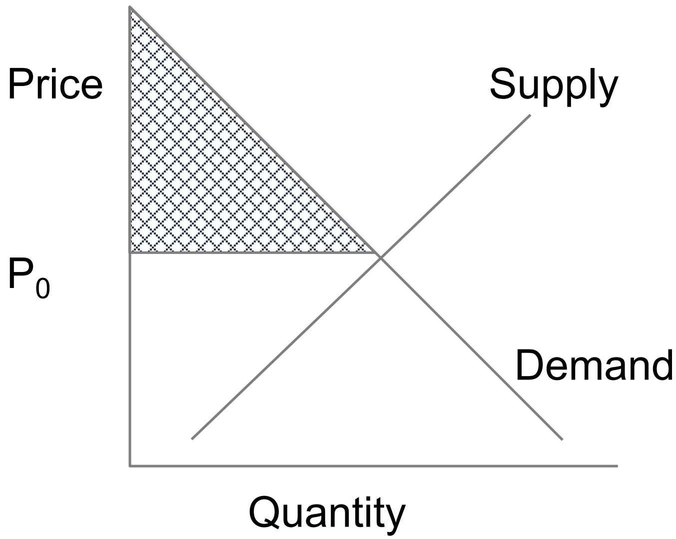
- Invented in the airline industry in the mid–1980s
- Well-suited to industries characterized by:
  - High fixed capital costs, e.g., airplanes, railroads, cruise ships, hotels
  - Low marginal costs, e.g., one more seat, one more berth, one more room
  - Commodity product that is difficult to differentiate from competitors
  - Competitive, dynamic pricing and demand *elasticity* as a function of price changes
- RM’s purpose is to *sell the right capacity to the right customer at the right price; setting differential prices for homogeneous capacity*
- Typically results in a three to seven percent increase in revenue

Revenue Management Overview

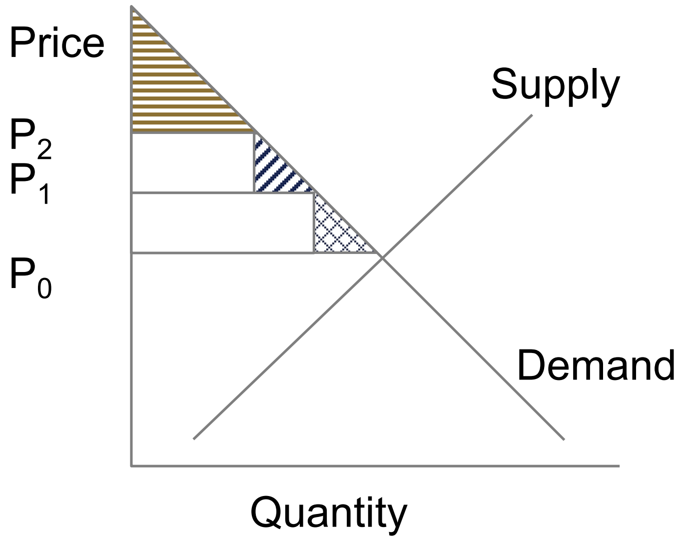
- Involves several problems that are solved using predictive and prescriptive techniques
- Customer segmentation
- Fare class segmentation based on product utility differentiation
- Demand forecasting by product subject to seasonality
- Inventory allocation optimization by product subject to demand, pricing

Revenue Management Overview

Pricing without revenue management



Pricing with revenue management



DataScience@SMU

Harrah’s Cherokee | Pachinko

Macro

Doug Gray

Harrah’s Cherokee Casino and Hotel

Harrah's Cherokee Casino and Hotel

- Located in Cherokee, NC
- Owned and operated by eastern band of the Cherokee Indians
- Managed by Harrah's (Total Rewards)
- 576 rooms
- No alcohol served
- Average room rate is $6/night
- Less than three hour drive from Atlanta, SC, NC, TN metro areas
- No direct competitors
- 98.6% occupancy rate January to November
- 60% margin on revenue
- RM yields 15% more revenue

Mr. Smith Tries to Book a Room…

- *Why is Mr. Smith's request to book a room rejected even though 183 of the 576 rooms are still unreserved? But he is comped a room at the nearby Ramada?*
  - Bets $2,000 per night netting the casino $140, probabilistically speaking

Mr. Smith Tries to Book a Room…

Problem

How does a casino hotel maximize revenue and subsequently maximize profitability?

- *Why is Mr. Smith's request to book at room rejected even though 183 of the 576 rooms are still unreserved? But he is comped a room at the nearby Ramada?*
  - Bets $2,000 per night netting the casino $140, probabilistically speaking

Mr. Smith Tries to Book a Room…

Problem

How does a casino hotel maximize revenue and subsequently maximize profitability?

- *Why is Mr. Smith's request to book at room rejected even though 183 of the 576 rooms are still unreserved? But he is comped a room at the nearby Ramada?*
  - Bets $2,000 per night netting the casino $140, probabilistically speaking

If ***N*** rooms are available, how can a system be devised to say "yes" to the ***N*** most profitable customers that wish to book on a given night, while declining the bookings of all other customers?

Mr. Smith Tries to Book a Room…

How does a casino hotel maximize revenue and subsequently maximize profitability?

- *Why is Mr. Smith's request to book at room rejected even though 183 of the 576 rooms are still unreserved? But he is comped a room at the nearby Ramada?*
  - Bets $2,000 per night netting the casino $140, probabilistically speaking

Problem

Customer segmentation

Demand forecast by segment

Willingness to pay (WTP) (or gamble)

Allocate capacity using bid price

Dynamic inventory capacity management

Maximize revenue by holding out and allocating hotel casino room capacity based on ranked WTP customer segments

Mr. Smith Tries to Book a Room…

How does a casino hotel maximize revenue and subsequently maximize profitability?

Problem

Descriptive

Predictive

Prescriptive

Customer segmentation

Demand forecast by segment

Willingness to pay (WTP) (or gamble)

Allocate capacity using bid price

Dynamic inventory capacity management

Maximize revenue by holding out and allocating hotel casino room capacity based on ranked WTP customer segments

Casino Hotel Revenue Management

Customer segmentation

Willingness to pay (or gamble)

Segment	Expected wagering profit
CS0	≥1,000
CS1	800–999
CS2	600–799
CS3	400–599
CS4	300–399
CS5	200–299
CS6	100–199
CS7	50–99
CS8	0–50
CS9	unknown
Total	

Mr. Smith Tries to Book a Room…

How does a casino hotel maximize revenue and subsequently maximize profitability?

Problem

If ***N*** rooms are available, how can a system be devised to say “yes” to the ***N*** most profitable customers that wish to book on a given night, while declining the bookings of all other customers?

Descriptive

Predictive

Prescriptive

Customer segmentation

Demand forecast by segment

Maximize revenue by holding out and allocating hotel casino room capacity based on ranked WTP customer segments

Willingness to pay (WTP) (or gamble)

Allocate capacity using bid price

Dynamic inventory capacity management

DataScience@SMU

Harrah’s Cherokee | Pachinko

Micro

Doug Gray

Casino Hotel Revenue Management

- Cherokee utilizes **Total Rewards** Harrah's customer loyalty program and transaction history system
- **Total Rewards** racks customer play including games, money wagered, time played
  - 83% of money wagered is "tracked play"
- Revenue means total customer spend, including gambling, food, and hotel room
  - Average revenue per customer is $565/night
  - "Comps" include reduced price rooms, casino chips, valet service, meals, and other perks and amenities
- Casinos and their customers incentives are *perfectly aligned; gamble more, earn more points and perks, generate more profit*

Descriptive

Customer segmentation

Willingness to pay (or gamble)

Customer Segment Demand Forecasting

- Unconstrained demand by customer segment
  - Total unconstrained demand = bookings made + booking requests denied
- Any modern day machine learning method can be employed to forecast demand
- Historically, *time series* methods, such as *exponential smoothing* provided an interpretable, accurate approach
  - Holt-Winters smoothing model was utilized in this case

Predictive

Demand forecast by segment

Customer Segment Demand Forecasting

- Holt-Winters forecasts values for several demand trends
  - Base demand and demand trends
  - Seasonality, e.g., annual, month, day of week
  - Special events
  - Easily interpretable and reasonably accurate
- Forecast demand for every customer segment, every day in the booking curve

Predictive

Demand forecast by segment

$$\textbf{\textit{Forecast}}_{i,t} = (\textbf{\textit{Base}}_{i,t} + \textbf{\textit{Trend}}_{i,t})(\textbf{\textit{Seasonal Month}}_{i,t})(\textbf{\textit{Seasonal Day}}_{i,t})(\textbf{\textit{Event}}_{i,t})$$

Capacity Allocation Optimization: Bid Price

- *Opportunity cost* of selling a unit of capacity to the wrong customer, i.e., one that will generate less revenue than another customer; the *marginal value* of one more unit of capacity
- *Average revenue* of higher value customer segments multiplied by the probabilities that a higher value customer will request that unit of capacity in the future
- *Threshold limit* for the amount of expected revenue a customer will generate per night

If *bid price > e(revenue)* then *reject reservation request*
If *bid price ≤ e(revenue)* then *accept reservation request*

Prescriptive

Dynamic inventory capacity management

Maximize revenue by holding out and allocating hotel casino room capacity based on ranked WTP customer segments

Capacity Allocation Optimization: LP Model

To demonstrate how a bid price control system works, assume for simplicity that a casino only rents rooms for one night stays (this procedure can easily be modified to accommodate multiple night stays but the notation becomes more complex) Also assume the casino segments its customers into K distinct segments. Let r_i represent the historical average revenue that a customer from segment i spends per night at the casino, $i = 1,\ldots K$. Let d_i represent the forecasted total demand for rooms from customers in segment i. The decision variable is X_i, the number of rooms to allocate to customers in segment i. The sum of the rooms allocated to each segment must be less than or equal to the capacity of the casino. Using this notation, the bid price control problem can be formulated as a linear program (LP) as follows:

$$\text{Max} \quad \text{Total revenue} = \sum_{i-l}^{K} r_i X_i$$

$$\textit{s.t.} \quad X_i \le d_i, \quad i = 1, .. K$$

$$X_i \ge 0, \quad i = 1, ..K$$

$$\sum_{i-l}^{K} X_i \le \text{capacity}$$

- Bid price is the “shadow price” of one additional unity of capacity
- Dual variable on the capacity constraint

Prescriptive

Dynamic inventory capacity management

Maximize revenue by holding out and allocating hotel casino room capacity based on ranked WTP customer segments

Demand forecast by segment

Dynamic inventory capacity management

Allocate capacity using bid price

Table 1: Cherokee customer segmentation scheme (approximate data):

Thursday data for a Friday event

Segment	Expected wagering profit	Unconstrained demand	Demand override	Rooms allocated	Current sold	Bid price
CS0	≥1,000	119	119	120	84	REB1*
CS1	800-999	128	128	122	75	REB2*
CS2	600-799	126	126	124	69	ROC*
CS3	400-599	122	**150**	138	79	ROC*
CS4	300-399	155	155	43	43	$125
CS5	200-299	168	168	0	0	$225
CS6	100-199	144	144	0	0	$325
CS7	50-99	103	103	0	0	$375
CS8	0-50	92	92	0	0	$425
CS9	unknown	45	45	0	0	$450

*RFB1 refers to “complimentary room, food, and beverage at level 1, the best rooms and restaurants available. RFB2 refers to a lesser level of complimentary meals. ROC indicates the “room only” is complimentary.

Maximize Revenue by Holding Out and Allocating Hotel Casino Room Capacity Based on Ranked WTP Customer Segments

Table 2: Possible revenue streams: representative weekend night

Segment	Expected wagering profit	Room sold: RM system	Gross revenue: RM system*	Rooms sold: optimal*	Gross revenue: optimal*	Room sold: no system	Gross revenue: no system**
CS0	≥1,000	120	$180,000	125	$187,500	55	$89,375
CS1	800–999	122	$109,800	130	$117,000	55	$56,375
CS2	600–799	124	$86,800	124	$86,800	55	$45,375
CS3	400–599	138	$69,000	140	$70,000	55	$34,375
CS4	300–399	43	$15,050	28	$9,800	55	$26,125
CS5	200–299	0	0	0	0	55	$20,625
CS6	100–199	0	0	0	0	55	$15,125
CS7	50–99	0	0	0	0	54	$10,800
CS8	0–50	0	0	0	0	54	$8,100
CS9	unknown	0	0	0	0	54	$12,150
Total		547	$460,650	547	$471,100	547	$318,425

* Gross revenue calculation assumes $1,5000/night for CSO segment, and the midpoint of the expected wagering profit for other segments.
** Gross revenue also assumes no complimentary rooms, a room rate of $125/night and $100 gaming revenue for segment CS9 (unknown customers).

Maximize Revenue by Holding Out and Allocating Hotel Casino Room Capacity Based on Ranked WTP Customer Segments

Table 2: Possible revenue streams: representative weekend night

Segment	Expected wagering profit	Room sold: RM system	Gross revenue: RM system*	Rooms sold: optimal	Gross revenue: optimal*	Room sold: no system	Gross revenue: no system**
CS0	≥1,000	120	$180,000	125	$187,500	55	$89,375
CS1	800–999	122	$109,800	130	$117,000	55	$56,375
CS2	600–799	124	$86,800	124	$86,800	55	$45,375
CS3	400–599	138	$69,000	140	$70,000	55	$34,375
CS4	300–399	43	$15,050	28	$9,800	55	$26,125
CS5	200–299	0	0	0	0	55	$20,625
CS6	100–199	0	0	0	0	55	$15,125
CS7	50–99	0	0	0	0	54	$10,800
CS8	0–50	0	0	0	0	54	$8,100
CS9	unknown	0	0	0	0	54	$12,150
Total		547	$460,650	547	$471,100	547	$318,425

Revenue management results in a ***45% increase in revenue*** over a *first-come-first-served naïve* allocation of rooms to customer reservation requests; achieves a revenue target that is *97.8%% of Optimal revenue; $142,000 in pure profit in one day*

(*Optimal revenue* is calculated in hindsight with perfect information)

Harrah's Cherokee Casino and Hotel RM

- Combination of ***descriptive, predictive, and prescriptive*** analytics is utilized to enable the revenue management system solution.
- Revenue management results in a ***15% average increase in total revenue.***
- Customers segment themselves with their gambling wagering levels.
- RM is integrated with CRM to target and optimize marketing initiatives; response rates are tracked and measured to determine which incentives work best.
- Initially, RM drives significant change throughout an organization in terms of people's task focus and span of control, as the system drives decision-making.

http://www.bloomberg.com/news/articles/2010-08-06/loveman-plays-new-purely-empirical-game-as-harrah-s-ceo

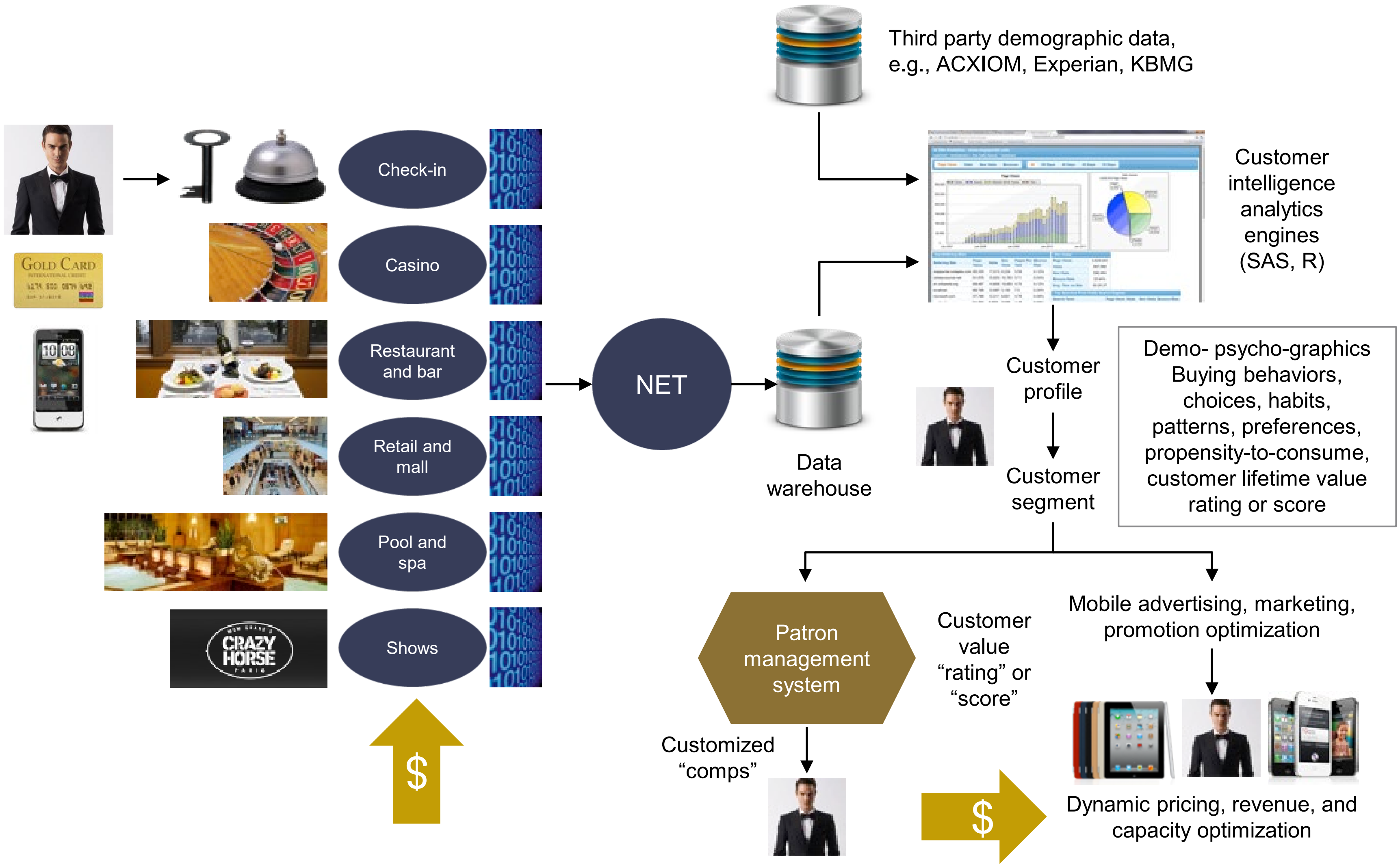
DataScience@SMU

Harrah's Cherokee

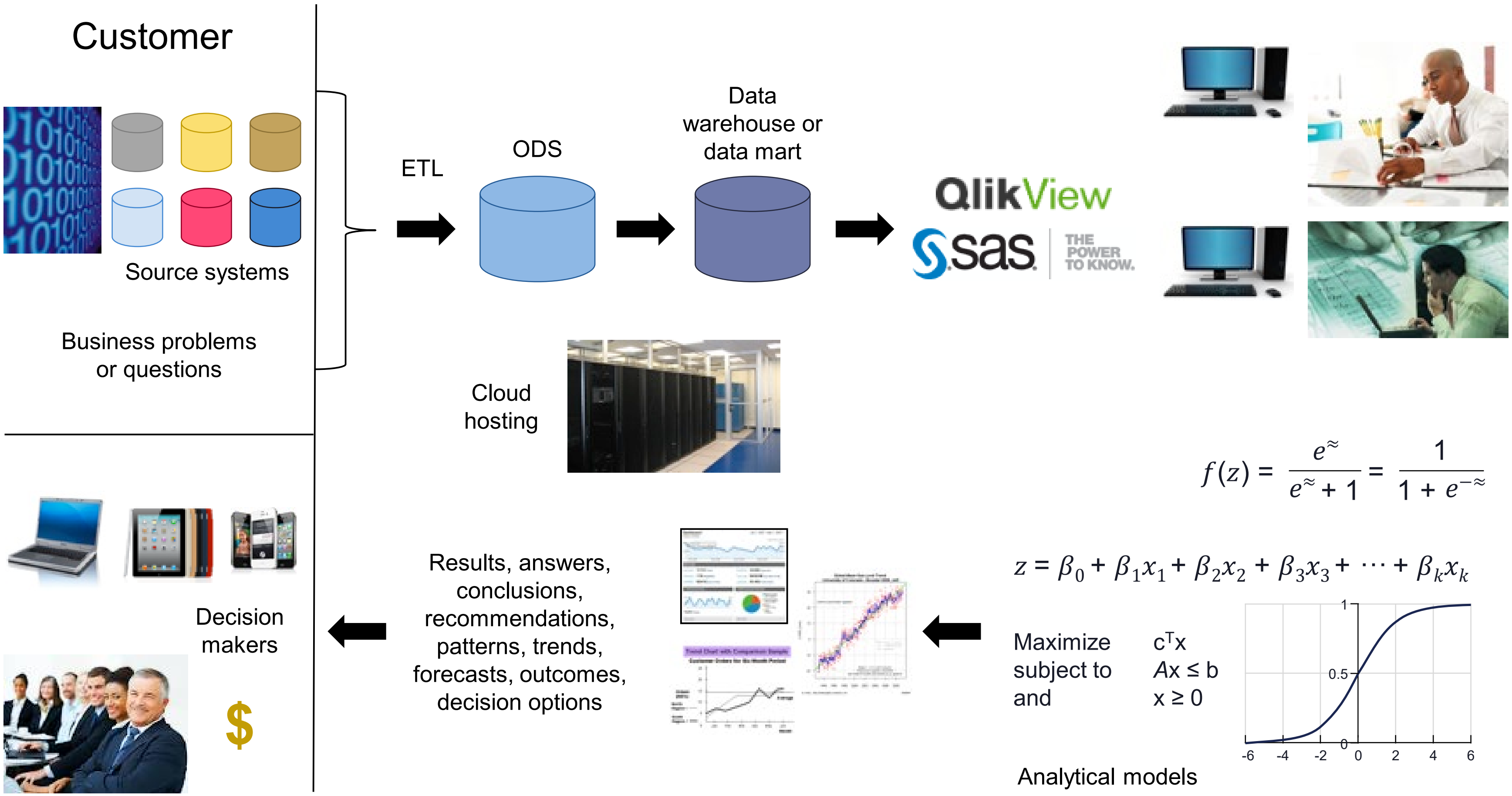
Analytics Architecture

Doug Gray

Analytics Architecture



Analytics Architecture Flow



DataScience@SMU

Bonus on Customer Segmentation

Bayesian Belief Networks

Doug Gray

Bonus on Customer Segmentation

- Predicting customer response to a gamification offer using Bayesian Belief Networks
- Intercontinental Hotels Group (IHG)

Alex Cosmas: Chief Scientist

Alex Cosmas is an Expert Associate Partner at McKinsey & Co., specializing in predictive analytics across the transportation, travel, and consumer sectors. He is a recognized expert in the use of probabilistic and causal models to perform both deductive and inductive reasoning from large datasets. He has consulted for Fortune 100's both domestically and internationally in the areas of demand modeling, consumer choice, network modeling, revenue management and pricing. He earned his B.S.in Applied Physics from Columbia Engineering and M.S. degrees in Technology Policy and Aerospace Engineering, both from the Massachusetts Institute of Technology.

http://www.bayesia.com/cosmas-so-you-can-predict-the-future

Bayesian Inference Model

- The "Swiss Army Knife" of *modern* statistical predictive modeling

$$P(M|E) = P(M) \cdot \frac{P(E|M)}{P(E)}$$

Breast Cancer Screening Indicator (BCSI) Test				
BCSI Test Result	True Cancer State	Patient Has Cancer	Patient Has No Cancer	Total # Patients Tested
POS +		88	36	124
NEG -		18	37	55
Total		106	73	179
P(Patient Has Cancer/POS + Test Result) = 88 / 124 = 71%				
P(POS + Test Result/Patient Has No Cancer) = 36 /173 = 49%				

Modeling Propensity of an Outcome Given *Prior* Information and *Empirical* Information

- Disease State given a Test Result
- Heart attack or stroke
- Hospital readmission
- Consumer behavior, e.g., product purchase, churn

Bayesian Inference Model

- The "Swiss Army Knife" of *modern* statistical predictive modeling

Breast Cancer Screening Indicator (BCSI) Test				
BCSI Test Result	True Cancer State	Patient Has Cancer	Patient Has No Cancer	Total # Patients Tested
POS +		88	36	124
NEG -		18	37	55
Total		106	73	179
P(Patient Has Cancer/POS + Test Result) = 88 / 124 = 71%				
P(POS + Test Result/Patient Has No Cancer) = 36 /173 = 49%				

$$P(M|E) = P(M) \cdot \frac{P(E|M)}{P(E)}$$

** Check using the Bayes formula:*

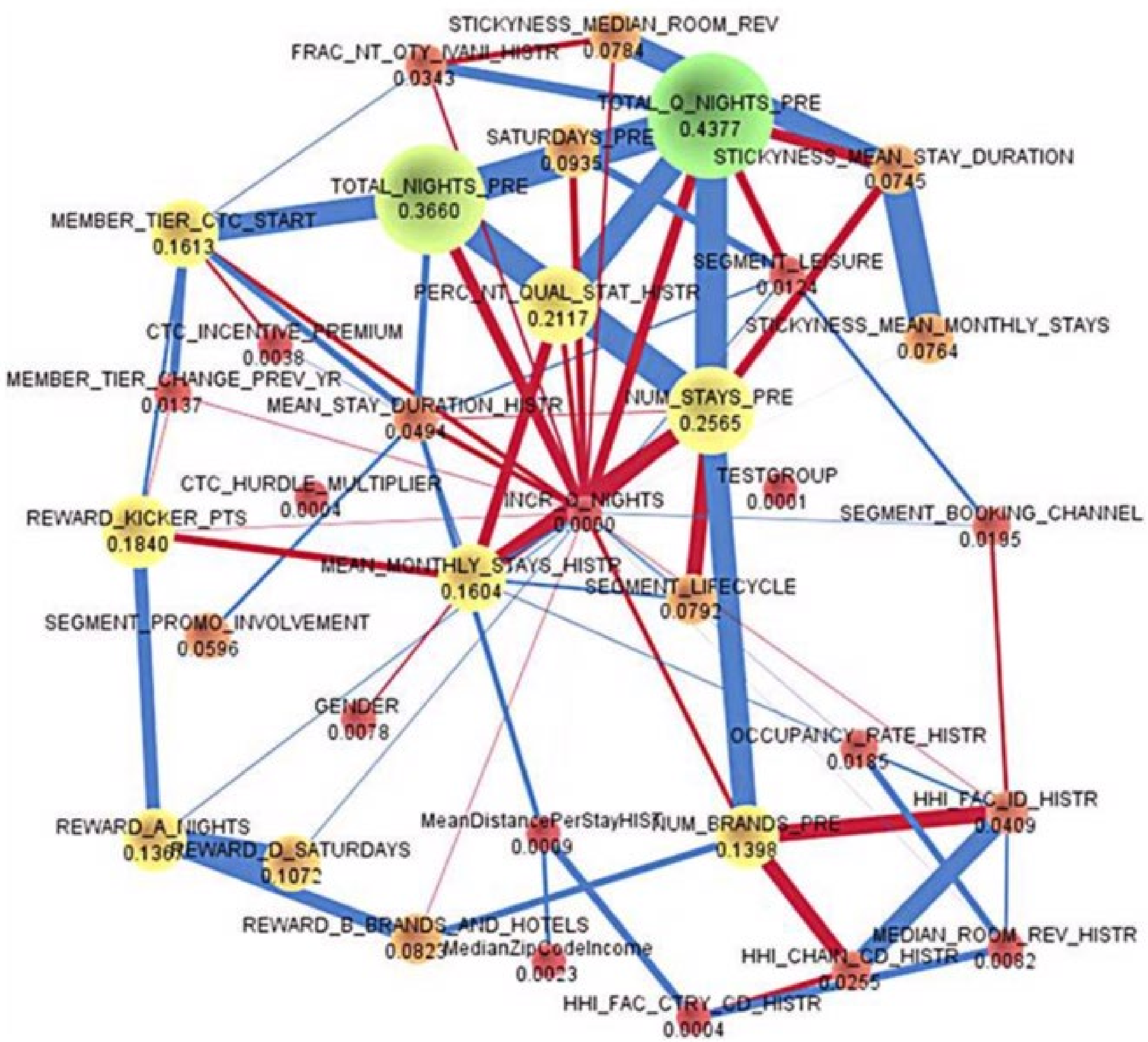
*P(C/+) = P(C) * P(+/C) / P(+)*

*= (106/179) * (88/106) / (124/179)*

= 88 / 124 = 71%

Model Performance

Figure 7 BBN with nodes sized by mutual information with the target node



Model Performance

Table III Linear regression *F*-values.

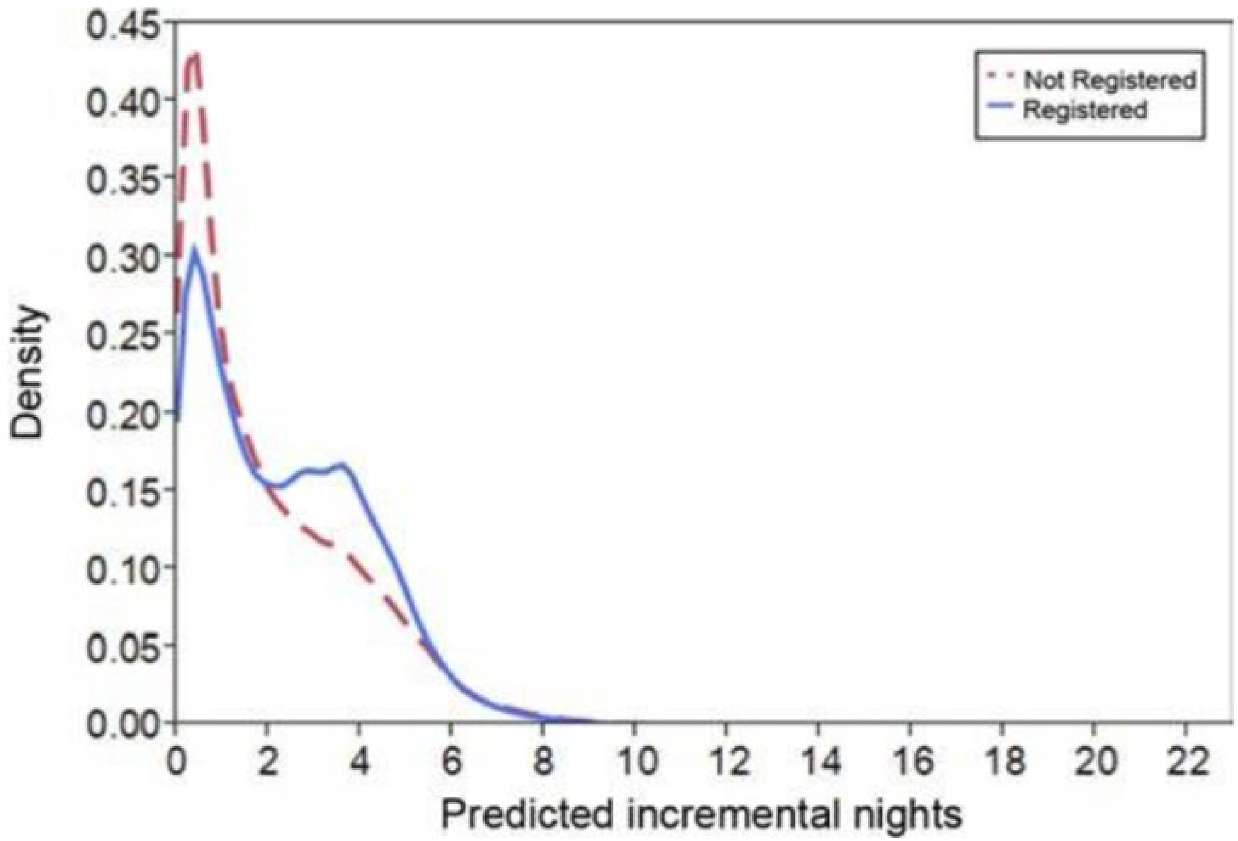
Variable	*F*-value
TOTAL_NIGHTS_PRE	33054.9
TOTAL_Q_NIGHTS_PRE	15062.85
NUM_STAYS_PRE	1922.855
NUM_BRANDS_PRE	375.5618

Table IV Parameter estimates from the BBN and linear regression

Variable	BBN parameter estimate	Linear regression parameter estimate
MeanDistancePerStayHISTR	0.0005	0.0006
OCCUPANCY_RATE_HISTR	2.0424	2.9843
MEDIAN_ROOM_REV_HISTR	0.0019	-0.0011

Model Performance

Figure 10 Kernel density plot for predict incremental nights by registration status



Top 10 Key Takeaways Regression vs. BBN

1. Estimating *Incremental Room Nights* resulting from participation in ***Crack the Case***
2. Use of *gamification* to increase customer engagement, (incremental) sales & brand awareness; net benefit of the marketing program
3. Comparing Bayesian Belief Networks (non-parametric) vs. Linear Regression (parametric)
4. BBN & LR made comparable predictions of TIRN driven by CTC (< 20% variance)
5. BBNs are non-parametric, require no control groups, make no assumption of variable independence, and use collinearity to derive the probability model and ID causal parent-child relationships

6. BBNs can identify secondary and tertiary relationships between and among the dependent variable and independent variables, to describe a relational hierarchy
7. BBN requires no exact look-alikes to estimate the influence of the promotional campaign
8. Opportunity to optimize campaign design; estimate promotional attributes that are likely to drive incremental spend
9. Operationalize more efficient audience selection
10. Clearly an opportunity for more data-driven approaches to optimize Marketing Campaign ROI

DataScience@SMU