



CARLSON SCHOOL
OF MANAGEMENT
UNIVERSITY OF MINNESOTA

ANALYTICS & CUSTOMER RESEARCH: MALL OF AMERICA GAINS INSIGHTS FROM WI-FI DATA

by Lee C. Thomas & Ravi Bapna

Analytics Teaching Case – Teaching Plan
Contact thom1221@umn.edu for material

Situation

- Online shopping and digital technologies are disrupting retail and the Mall culture – a staple of American life for generations



ONE OF THE LARGEST
FINANCIAL SERVICES
FIRMS IN THE COUNTRY.

LEARN MORE

Stores are closing at an epic pace

by Jackie Wattles @jackiewattles

April 22, 2017: 3:58 PM ET

Recommend



Social



Dying shopping malls are wreaking havoc on suburban America

Hayley Peterson
Mar. 5, 2017, 7:30 AM 134,343

Mall of America

- Planned as a shopping and entertainment **destination**
- Massive – 7 Yankee stadiums can fit inside
- Has roller coasters
- Aquarium
- Two hotels
- Nightclub
- 3 million international visitors
- 40% of 42 million visitors come from more than 150 miles
- *Needs to be at the edge of the consumer experience and has a desire to play in the digital sphere*

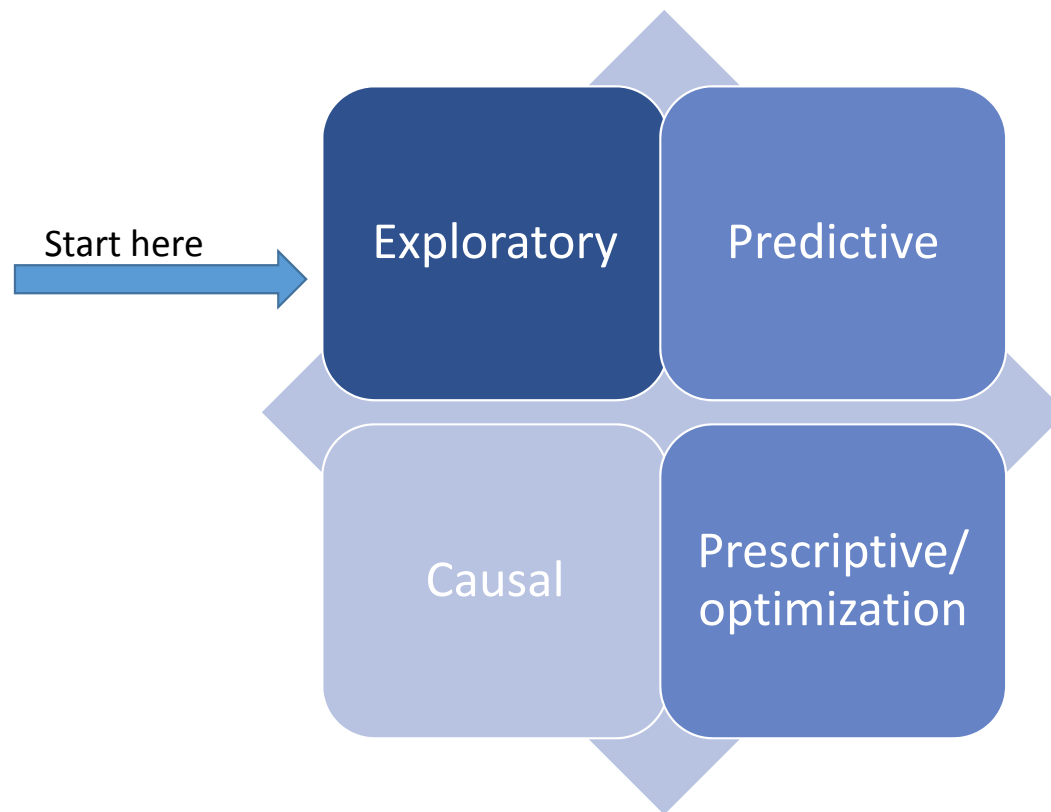
Complication

- Consumers expect free Wifi in spaces such as MOA
- But, it's a huge (\$7 million) investment and the business case has to be beyond "consumer's expect it"
- Data is the new oil is a commonly heard phrase and analytics is hot
 - However, MOA has no organizational capabilities in machine learning or data engineering for analytics
 - There isn't an awareness of what is possible

Key questions

- Is there business value to the MOA in giving free Wifi to its consumers?
- What kinds of use-cases exist from the mining of wifi-data?
- What does a typical analytics project look like? What skill sets does MOA need for this?
- How to design a pilot study that serves as a proof-of-concept to show the business value individual level wifi usage data?
- Where could this go in the future?

Four pillars of analytics



Solution approach

- Run segmentation analysis to better understand patterns of users mall visiting behaviors
 - Dataset of 772,019 rows and multiple attributes relating to an individual mall visit are provided based on when and where people touch the wifi access points
- Students have to do basic data cleansing and engineering
 - Box-cox transformations
 - Missing values
 - Derived features
- Run an unsupervised machine learning algorithm (k-means) and determine optimal number of clusters using the elbow criteria
- Identify, name and describe clusters of mall visits
- Make meaningful recommendations to answer the use-cases identified earlier

Data dictionary

Below is the data dictionary for the MOA Analytics table:

Column	Definition
device	MAC address of the device
session_date	Date of unique visit (A unique visit includes all session updates in a time window from 4 AM
entry_time	Time of first session-update with mall wifi
exit_time	Time of last session-update with mall wifi
duration	Sum of duration of all session updates (in seconds)
avg_time_session_update	Duration / Total number of session updates (in seconds)
distinct_ap	Distinct APs connected to
total_updates	Total number of session updates
time_per_distinct_ap	Duration / Distinct APs connected to
total_data_in	Total data received by the device
total_data_out	Total data sent by the device
distinct_sections_visited	Total distinct level-sections visited
distinct_levels_visited	Total distinct levels visited
weekday_flag	Visit on weekday or weekend. 1 stands for weekend
day_of_week	Day of week(starts from Sunday, i.e. 1 = Sunday, 6 = Friday)
time_0_1	Total time spent in level 0.1
time_0	Total time spent in level 0
time_1	Total time spent in level 1
time_2	Total time spent in level 2
time_3	Total time spent in level 3
time_4	Total time spent in level 4
time_0_E	Total time spent in section 0 East
time_0_N	Total time spent in section 0 North
time_0_NE	Total time spent in section 0 Northeast
time_0_S	Total time spent in section 0 South
time_0_SE	Total time spent in section 0 Southeast
time_1_E	Total time spent in section 1 East
time_1_N	Total time spent in section 1 North
time_1_NE	Total time spent in section 1 Northeast
time_1_NEx	Total time spent in section 1 North Extension
time_1_NW	Total time spent in section 1 Northwest
time_1_P	Total time spent in section 1 Park
time_1_S	Total time spent in section 1 South
time_1_SE	Total time spent in section 1 Southeast
time_1_SW	Total time spent in section 1 Southwest
time_1_W	Total time spent in section 1 West
time_2_E	Total time spent in section 2 East
time_2_N	Total time spent in section 2 North
time_2_NE	Total time spent in section 2 Northeast

Instructors resources

- Case narrative
- Datasets of wifi logs
- Technical note (20 pages) in R to run data engineering and k-means
 - Sample description of clusters