CUSTOMER ANALYTICS

LECTURE 1: INTRODUCTION

ANA ALINA TUDORAN

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

JANUARY - FEBRUARY 2024

ANA ALINA TUDORAN, PHD
ASSOCIATE PROFESSOR

AGENDA

I. Lecturers
II. Plan and teaching materials
III. Course organization
IV. Exam and exam questions
V. Overview of customer analytics methods

LECTURERS

Morten Berg Jensen (MBJ)

Expertise: PLS, segmentation, recommender systems

Style: technical and methodical style

Ana Alina Tudoran (AAT)

Expertise: SEM, Bayesian networks, Web analytics

Style: experiential and instrumental style

COURSES WITH +2 LECTURERS

- Each of us brings a valuable perspective
- We believe you can benefit greatly by learning from both of us
- Instead of comparing or rating us, we encourage you to appreciate the diversity of knowledge and skills we offer
- Our goal is to help you develop a well-rounded understanding of the subject by drawing from our complementary strengths

COURSE DETAILS

Highlights:

- 10 ECTS course
- 2X2 lectures per week in 14 weeks
- Prerequisites:
    - Machine Learning for Business Intelligence I, or a similar
    - R programming

QUESTIONS

- If you send us e-mails, use your AU email. Address the question to the lecturer responsible for the topic.

- Address all the administrative questions to **Gitte Isager** at gi@econ.au.dk

LECTURE PLAN

- 4 topics
- 5 main methods
- They are treated in depth

Calendar week (order)	Topic	Lecturer
	INTRODUCTION	
5 (1)	Course Organization & Objectives	AAT
	EXPLAINING CUSTOMER BEHAVIOR	
5 (2)	Factor Analysis (I)	AAT
6 (3)	Factor Analysis (II)	AAT
6 (4)	Structural Equation Modelling (I)	AAT
7 (5)	Structural Equation Modelling (II)	AAT
7 (6)	Structural Equation Modelling (III)	AAT
8 (7)	*Application/Case Study*	AAT
8 (8)	Partial Least Squares (I)	MBJ
9 (9)	Partial Least Squares (II)	MBJ
9 (10)	Partial Least Squares (III)	MBJ
10 (11)	Partial Least Squares (IV)	MBJ
	CAUSAL DISCOVERY, PRODUCT RECOMMENDATION & TARGETING (I)	
10 (12)	Bayesian Networks (I)	AAT
11 (13)	Bayesian Networks (II)	AAT
11 (14)	Bayesian Networks (III)	AAT
	WEB ANALYTICS	
12 (15)	Click-path analysis (I)	AAT
12 (16)	Click-path analysis (II)	AAT
14 (17)	Choice modeling (I)	AAT
14 (18)	Choice modelling (II) (1h) and *Description of Company Case Competition (1h)*	AAT/MBJ
	CUSTOMER SEGMENTATION	
15 (19)	Segmentation (I)	MBJ
15 (20)	Segmentation (II)	MBJ
16 (21)	Segmentation (III)	MBJ
16 (22)	Segmentation (IV)	MBJ
17 (23)	Segmentation (V)	MBJ
	PRODUCT RECOMMENDATION & CUSTOMER TARGETING (II)	
17 (24)	Association rules mining	MBJ
18 (25)	Collaborative filtering (CL) and Memory-based methods	MBJ
18 (26)	CL Model-based methods: Latent factor models and matrix factorization	MBJ
19 (27)	Wrap up product recommendation & customer targeting (II)	MBJ
19 (28)	*Students' presentations, incl. company feedback & reward for the best solutions (2h)*	AAT/MBJ

AARHUS BSS

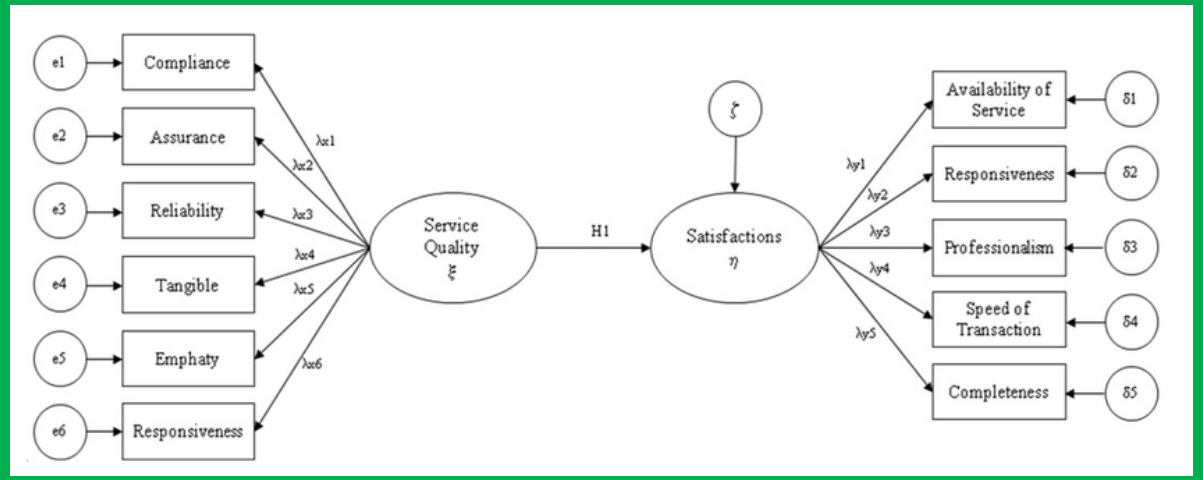
EXPLAIN BEHAVIOR

PREDICT CHOICES & BEHAVIOR

SEGMENT CUSTOMERS

PREDICT RATINGS TO RECOMMEND PRODUCTS

EXPLAIN BEHAVIOR, E.G.



PREDICT CHOICES & BEHAVIOR, E.G.

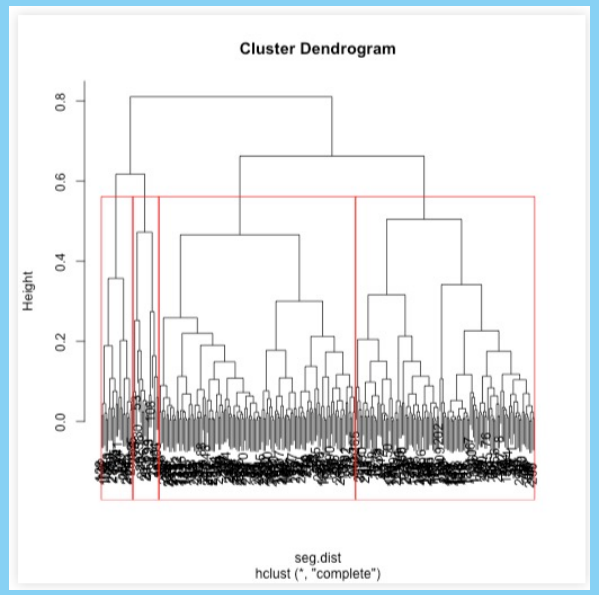
Which of the following minivans would you buy?
Assume all three minivans are identical other than the features listed below.

	Option 1	Option 2	Option 3
	6 passengers	8 passengers	6 passengers
	2 ft. cargo area	3 ft. cargo area	3 ft. cargo area
	gas engine	hybrid engine	gas engine
	$35,000	$30,000	$30,000
I prefer (check one):	☐	☐	☑

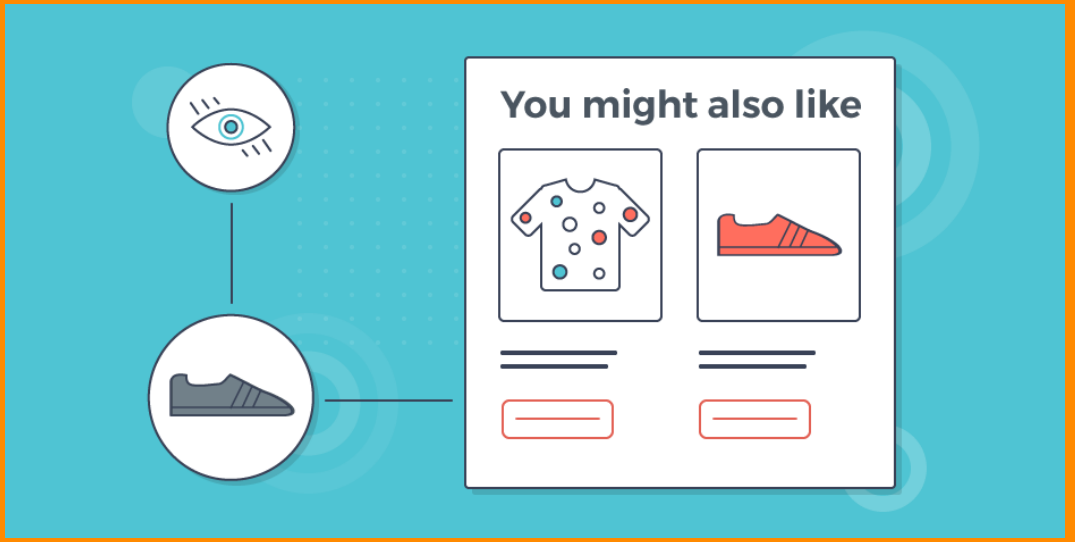
Fig. 13.1. An example choice-based conjoint survey question.

Respondents can answer up to 50 of these types of questions, which gives us lots of data from which to infer their preferences.

SEGMENT CUSTOMERS, E.G.



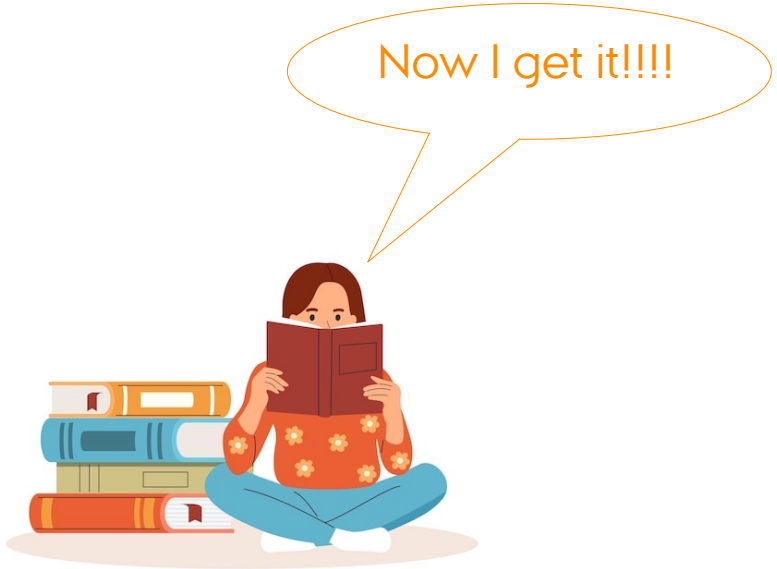
RECOMMEND PRODUCTS, E.G.



TEACHING MATERIAL

Notice that lectures include:

- Slides
- R files
- Videos
- Support material for each lecture



FOR NEXT LECTURE

Weeks 5(2) – 8(7). EXPLAINING CUSTOMER BEHAVIOR

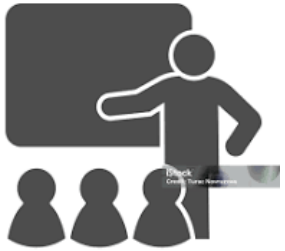
FOR FACTOR ANALYSIS:

- Mehmetoglu M. and Mittner M. (2022). *Applied Statistics using R: A guide for the Social Sciences*, Sage. **Chapter 13.** Available on Brightspace.

HOW TO USE THE MATERIAL

SUPPORT MATERIAL
Have a look before lecture

SLIDES & R CODE
Participate in the lecture. We present the concepts, the challenges, and hands-on exercises



SUPPORT MATERIAL
Read it carefully after the lecture/when preparing exam

SLIDES & R CODE
Rehearse again the applications (R code) when preparing exam

EXAM DETAILS

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

JANUARY - FEBRUARY 2024

ANA ALINA TUDORAN, PHD
ASSOCIATE PROFESSOR

EXAM FORMAT

EXAM ASSIGNMENT

Study Programme and level	**MSc Business Intelligence**						
Term	Summer 23o						
Course name and exam code(s)	Customer Analytics						460202E007
Exam form and duration	WOA: On-site written exam submitted digitally in WISEflow, use of the internet NOT allowed during the exam, own PC required.						4 hours
Date and time	6 June 2023						09.00-13.00
Supplementary material/aids	All	X	No		Specified		
Hand-in of hand-written material allowed	Yes		No	X			
Hand-in of extra material (appendix) in WISEflow allowed	Yes		No	X			
Anonymous exam	Yes	X	No		Comments: Please do **not** write your name or student ID number anywhere. Use your flow-id number (find this on the cover sheet).		
Other relevant information	**Avoid being suspected of exam cheating.** Remember to state references and use quotation marks, if you copy text from other sources or re-use parts of a previously submitted exam paper (plagiarism and self-plagiarism). Students must answer the exam assignment **individually**. All submitted exam papers are checked for plagiarism, so cheating and collaboration between students will be detected. A dataset is uploaded to WISEflow as appendix.						
Number of pages (incl. front page)	4 pages						

AARHUS BSS

EXAM FORMAT

EXAM ASSIGNMENT

Study Programme and level	**MSc Business Intelligence**						
Term	Summer 23o						
Course name and exam code(s)	Customer Analytics						460202E007
Exam form and duration	WOA: On-site written exam submitted digitally in WISEflow, use of the internet NOT allowed during the exam, own PC required.						4 hours
Date and time	6 June 2023						09.00-13.00
Supplementary material/aids	All	X	No		Specified		
Hand-in of hand-written material allowed	Yes		No	X			
Hand-in of extra material (appendix) in WISEflow allowed	Yes		No	X			
Anonymous exam	Yes	X	No		Comments: Please do **not** write your name or student ID number anywhere. Use your flow-id number (find this on the cover sheet).		
Other relevant information	**Avoid being suspected of exam cheating.** Remember to state references and use quotation marks, if you copy text from other sources or re-use parts of a previously submitted exam paper (plagiarism and self-plagiarism). Students must answer the exam assignment **individually**. All submitted exam papers are checked for plagiarism, so cheating and collaboration between students will be detected. A dataset is uploaded to WISEflow as appendix.						
Number of pages (incl. front page)	4 pages						

AARHUS BSS

EXAM FORMAT

EXAM ASSIGNMENT

Study Programme and level	**MSc Business Intelligence**						
Term	Summer 23o						
Course name and exam code(s)	Customer Analytics						460202E007
Exam form and duration	WOA: On-site written exam submitted digitally in WISEflow, use of the internet NOT allowed during the exam, own PC required.						4 hours
Date and time	6 June 2023						09.00-13.00
Supplementary material/aids	All	X	No		Specified		
Hand-in of hand-written material allowed	Yes		No	X			
Hand-in of extra material (appendix) in WISEflow allowed	Yes		No	X			
Anonymous exam	Yes	X	No		Comments: Please do **not** write your name or student ID number anywhere. Use your flow-id number (find this on the cover sheet).		
Other relevant information	**Avoid being suspected of exam cheating.** Remember to state references and use quotation marks, if you copy text from other sources or re-use parts of a previously submitted exam paper (plagiarism and self-plagiarism). Students must answer the exam assignment **individually**. All submitted exam papers are checked for plagiarism, so cheating and collaboration between students will be detected. A dataset is uploaded to WISEflow as appendix.						
Number of pages (incl. front page)	4 pages						

AARHUS BSS

EXAM CONTENT

- A set of questions **both theoretical and practical** covering the curriculum
- Exam and re-exam from 2022 and 2023 are available on Brightspace

EXAM PASSING CRITERIA

PART A. Lecturer: Ana Alina Tudoran: 50% pts

PART B. Lecturer: Morten Berg Jensen: 50% pts

Both sections must meet the minimum point requirement for a pass.

QUESTION (EXAM 2023)

The influencer model is described on pages 70-73 in your curriculum textbook (Hair Jr et al., 20211).(…). To solve this assignment, you must use the dataset in the file called "PLS_data_exam.csv".

1) Discuss briefly how to treat missing values in PLS analyses in general. (5 pts).

- This is a theoretical question.

2) Specify and run the influencer model.

- This question requires data analysis in R.

QUESTION (EXAM 2023)

Assuming a segmentation of the respondents based on the answers to variables “sic_1” to “sic_6” using cluster analysis:

3) Discuss considerations regarding the use of the three types of cluster analyses: classical cluster analysis, model-based cluster analysis, and latent class analysis.

- This question might seem theoretical.
- However, it is crucial to connect to the case study.
- Consider variables 'sic_1' to 'sic_6’ and how each clustering method could provide insights into customer segments in this case study. For instance, one can describe
    - How each method approaches the grouping of customers, in general.
    - What unique aspects of customer behavior or preference might each method reveal *in this case*?
    - How can the characteristics of 'sic_1' to 'sic_6' be best understood through each clustering technique?

SUBMISSION

In this examination, you must submit your full **answers in a Word/PDF document**. You should selectively copy and paste only the pertinent R code into Word and provide an analysis of the outcomes, as required in the exercise.

Q & A

INTRODUCTION TO CUSTOMER ANALYTICS

AARHUS BSS
DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

JANUARY - FEBRUARY 2024

ANA ALINA TUDORAN, PHD
ASSOCIATE PROFESSOR

CUSTOMER ANALYTICS

"The process of
collecting, analyzing, and interpreting
data about customers
to enhance business decision-making and performance"

COLLECTING DATA

	Data Sources	Use of each source
1	Transactional data	Records of customer purchases and spending patterns
2	CRM systems	Customer information, interactions & communication history
3	Web and e-commerce	Online behavior, clicks, and digital engagement metrics
4	Social media analytics	Brand sentiment, customer interests, and social media trends
5	Surveys & experiments	Direct customer feedback, attitudes, satisfaction, preferences
6	Loyalty programs	Customer loyalty insights and program effectiveness
7	Mobile data	App usage, mobile behavior, and location preferences
8	Third-Party data providers	External demographic and lifestyle data for context
9	IoT data	Real-time usage data from connected devices
10	Email & communication records	Analysis of communication effectiveness and customer queries

EXPLAIN BEHAVIOR

PREDICT CHOICES & BEHAVIOR

SEGMENT CUSTOMERS

PREDICT RATINGS TO RECOMMEND PRODUCTS

FACTORS EXPLAINING CUSTOMER LOYALTY

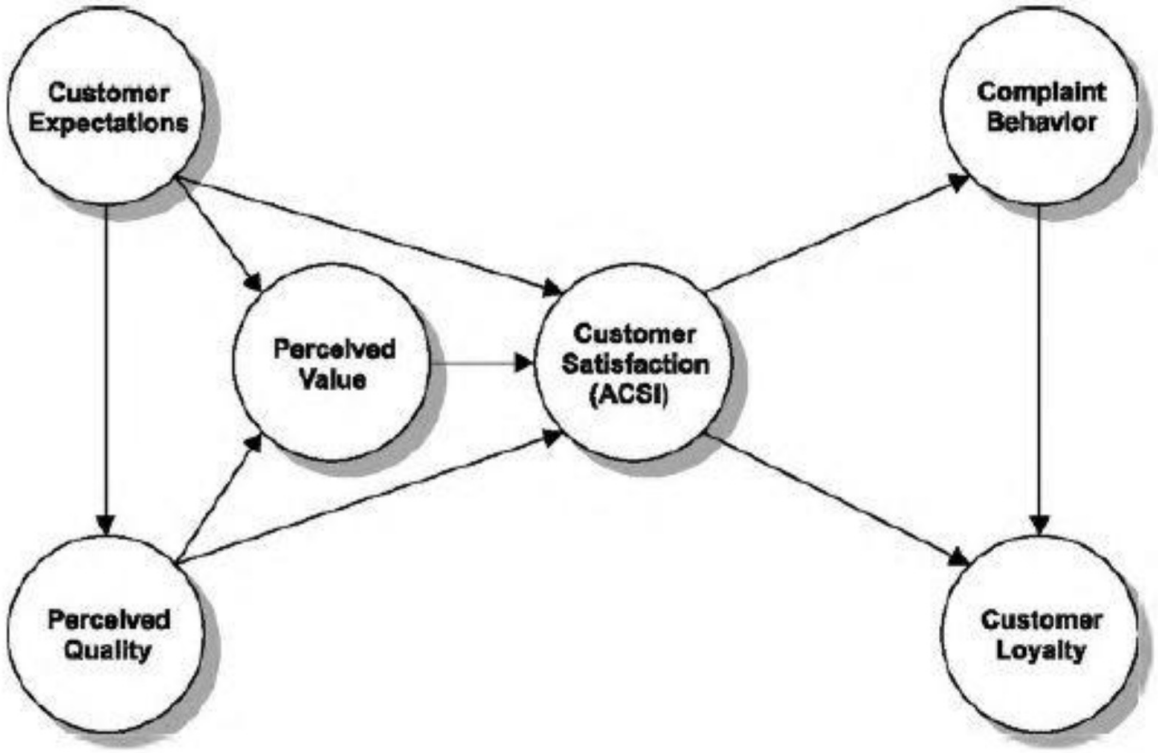


Fig. 2. The ACSI (American Customer Satisfaction Index) model.

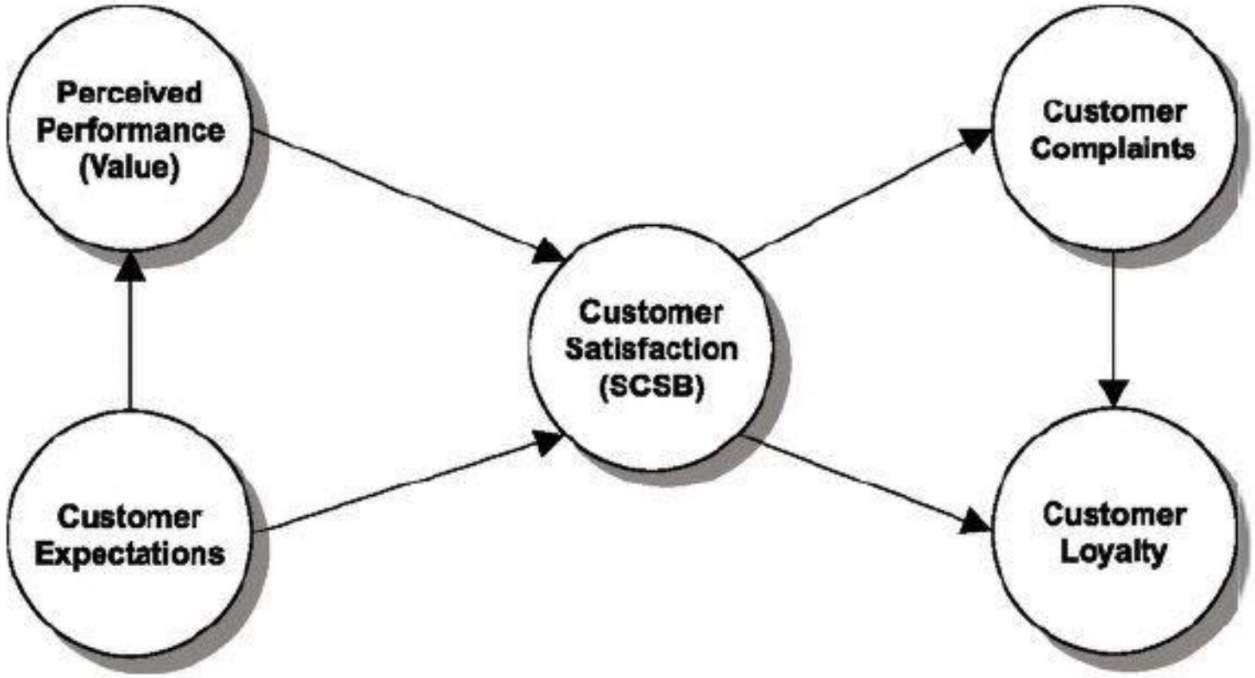


Fig. 1. The original SCSB (Swedish Customer Satisfaction Barometer) model.

Johnson, M. D., Gustafsson, A., Andreassen, T. W., Lervik, L., & Cha, J. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, 22(2), 217-245

SURVEY DATA

- Consider a standard satisfaction survey
  - Attitudes (agree/disagree)
  - Self-reported behavior
  - Demographics

- Each attitude statement belongs to a **factor,** latent variable or construct.

Measurement variable	Latent variable
1. Overall satisfaction	Customer satisfaction
2. Performance versus the customer's ideal service provider in the category	Customer satisfaction
3. Expectancy disconfirmation (performance that falls short of or exceeds expectations)	Customer satisfaction
4. Price compared to quality	Price
5. Price compared to other companies	Price
6. Price compared to expectations	Price
7. Corporate image compared to other companies	Corporate image
8. Image of the store (branch) you deal with	Corporate image
9. What friends say about the corporate image	Corporate image
10. Overall corporate image	Corporate image
11. The compensation offered by the company	Complaint behavior
12. Employees treated you politely and with respect when you complained	Complaint behavior
13. The pleasure taken in being a customer of the company	Affective commitment
14. Identification with what the company stands for	Affective commitment
15. Presence of reciprocity in the relationship	Affective commitment
16. Feeling of belongingness to the company	Affective commitment
17. The economics (benefits versus costs) of the alternative	Calculative commitment
18. Economic suffering if the relationship is broken	Calculative commitment
19. Location advantages versus other companies	Calculative commitment
20. Likelihood of retention	Loyalty
21. Likelihood of speaking favorably about the company to others	Loyalty
22. Likelihood of recommending the company to others	Loyalty

MODEL DEVELOPMENT & TESTING

(FACTOR ANALYSIS, SEM AND PLS-SEM)

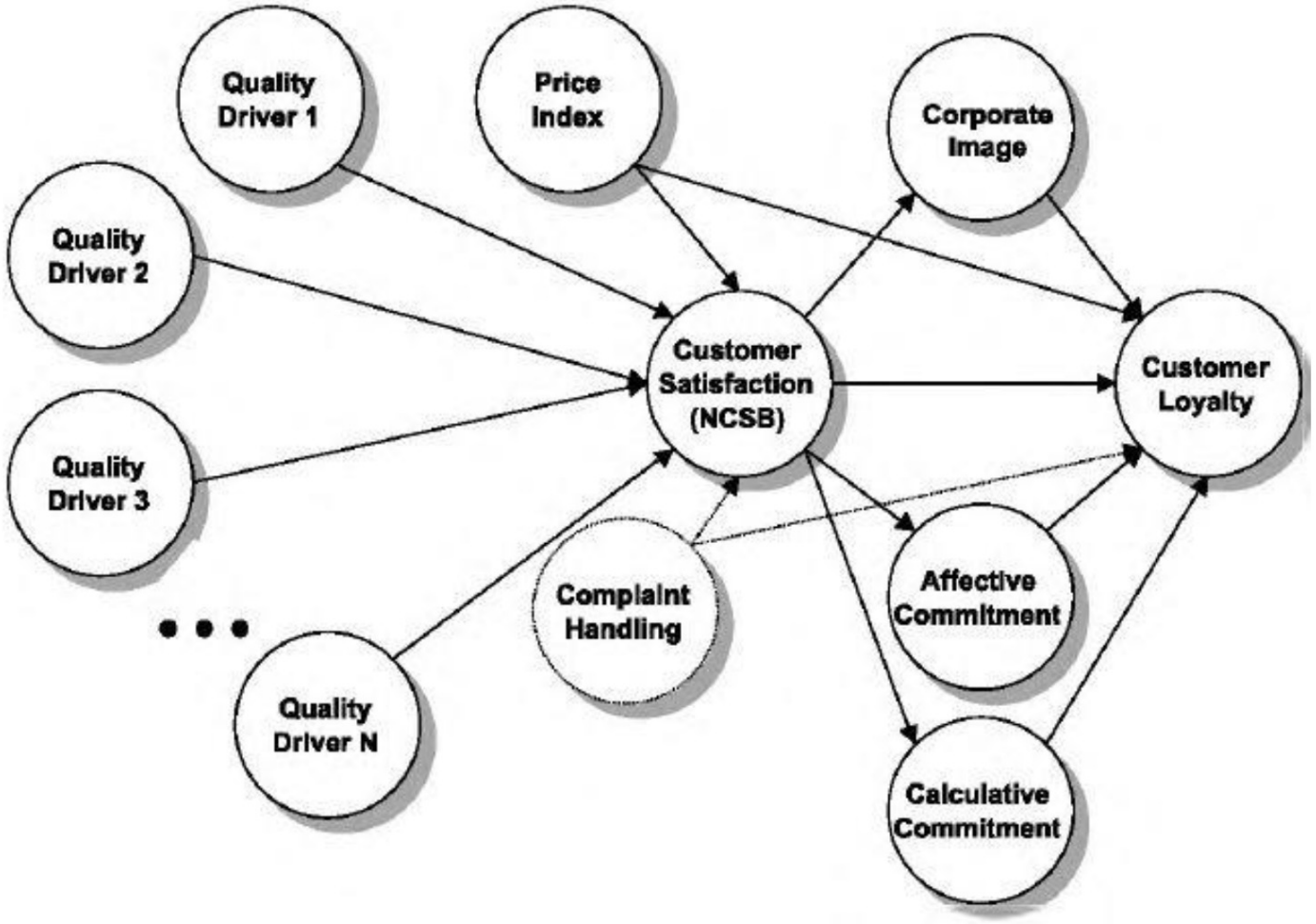


Fig. 3. The proposed model.

Johnson, M. D., Gustafsson, A., Andreassen, T. W., Lervik, L., & Cha, J. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, 22(2), 217-245

AARHUS BSS

COMPARATIVE OVERVIEW

Concept	Factor Analysis (FA)	Structural Equation Modeling (SEM)	Partial Least Squares SEM (PLS-SEM)
Primary Use	To identify underlying **factors** or latent variables that explain the patterns of correlations within a set of observed variables (items).	To test complex **cause-effect relationship** models with latent variables.	To analyze complex cause-effect relationship models with latent variables, particularly **when the assumptions of SEM are not met.**
Data Requirements	Large datasets; assumes variables (items) are **metric** and **normally distributed**.	Large sample sizes; assumes **multivariate normality** and **linearity**.	**Smaller sample sizes** are acceptable; **less stringent** on data distribution assumptions.

COMPARATIVE OVERVIEW

Concept	Factor Analysis (FA)	Structural Equation Modeling (SEM)	Partial Least Squares SEM (PLS-SEM)
Statistical Approach	**Reduces dimensionality** of data	Combines **factor analysis and multiple regression** analysis; uses a covariance-based approach.	Focuses on maximizing **explained variance of DVs**.
Model Characteristics	**Explores** data to find patterns; typically used in exploratory research.	**Confirms theoretical models**; used in confirmatory research.	Flexibility in terms of data requirements; used **when prediction is a goal.**
Complexity	Less complex; to **understand basic structure in data.**	**More complex;** requires understanding of both factor analysis and regression.	**More complex;** balances between exploratory and confirmatory approaches.

Q & A

EXPLAIN BEHAVIOR

PREDICT CHOICES & BEHAVIOR

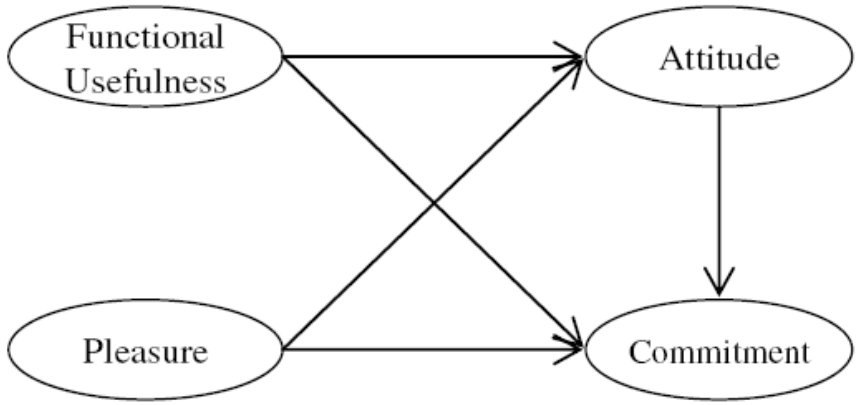
SEGMENT CUSTOMERS

PREDICT RATINGS TO RECOMMEND PRODUCTS

PREDICTING CUSTOMER COMMITMENT

Functional Usefulness		
p(Low)	*p*(Med)	*p*(High)
0.02	0.26	0.72

FUSE	PLEA	Attitude		
		p (Low)	*p* (Med)	*p* (High)
Low	Low	0.99	0.00	0.00
Low	Med	0.00	0.67	0.33
Low	High	0.00	0.99	0.00
Med	Low	0.33	0.33	0.33
Med	Med	0.00	0.79	0.21
Med	High	0.00	0.40	0.60
High	Low	0.99	0.00	0.00
High	Med	0.00	0.47	0.53
High	High	0.00	0.09	0.91



Pleasure		
p(Low)	*p*(Med)	*p*(High)
0.01	0.55	0.44

FUSE	PLEA	ATTI	Commitment		
			p (Low)	*p* (Med)	*p* (High)
Low	Low	Low	0.00	1.00	0.00
Low	Low	Med	0.33	0.33	0.33
Low	Low	High	0.33	0.33	0.33
Low	Med	Low	0.33	0.33	0.33
Low	Med	Med	0.00	1.00	0.00
Low	Med	High	1.00	0.00	0.00
Low	High	Low	0.33	0.33	0.33
Low	High	Med	0.00	1.00	0.00
Low	High	High	0.33	0.33	0.33
Med	Low	Low	0.33	0.33	0.33
Med	Low	Med	0.33	0.33	0.33
Med	Low	High	0.33	0.33	0.33
Med	Med	Low	0.33	0.33	0.33
Med	Med	Med	0.00	0.98	0.02
Med	Med	High	0.00	0.83	0.17
Med	High	Low	0.33	0.33	0.33
Med	High	Med	0.00	0.33	0.67
Med	High	High	0.00	0.44	0.56
High	Low	Low	1.00	0.00	0.00
High	Low	Med	0.33	0.33	0.33
High	Low	High	0.33	0.33	0.33
High	Med	Low	0.33	0.33	0.33
High	Med	Med	0.00	0.84	0.16
High	Med	High	0.00	0.71	0.29
High	High	Low	0.33	0.33	0.33
High	High	Med	0.00	0.40	0.60
High	High	High	0.00	0.10	0.90

Table 5
Forward inference due to change in different states of functional usefulness

State	Variables					
	FUSE		ATTI		COMM	
	PCP	NCP	PCP	NCP	PCP	NCP
Low	**0.02**	**1.00**	0.01	0.01	0.01	0.18
Medium	0.26	0.00	0.39	0.81	0.54	0.81
High	0.72	0.00	0.60	0.18	0.45	0.00

ATTI: attitude, FUSE: functional usefulness, COMM: commitment, PCP: prior conditional probability, NCP: new conditional probability.

Gupta, S. and Kim, H.W. (2008). Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities, European Journal of Operational Research, 190, 818-833

COMPARATIVE OVERVIEW

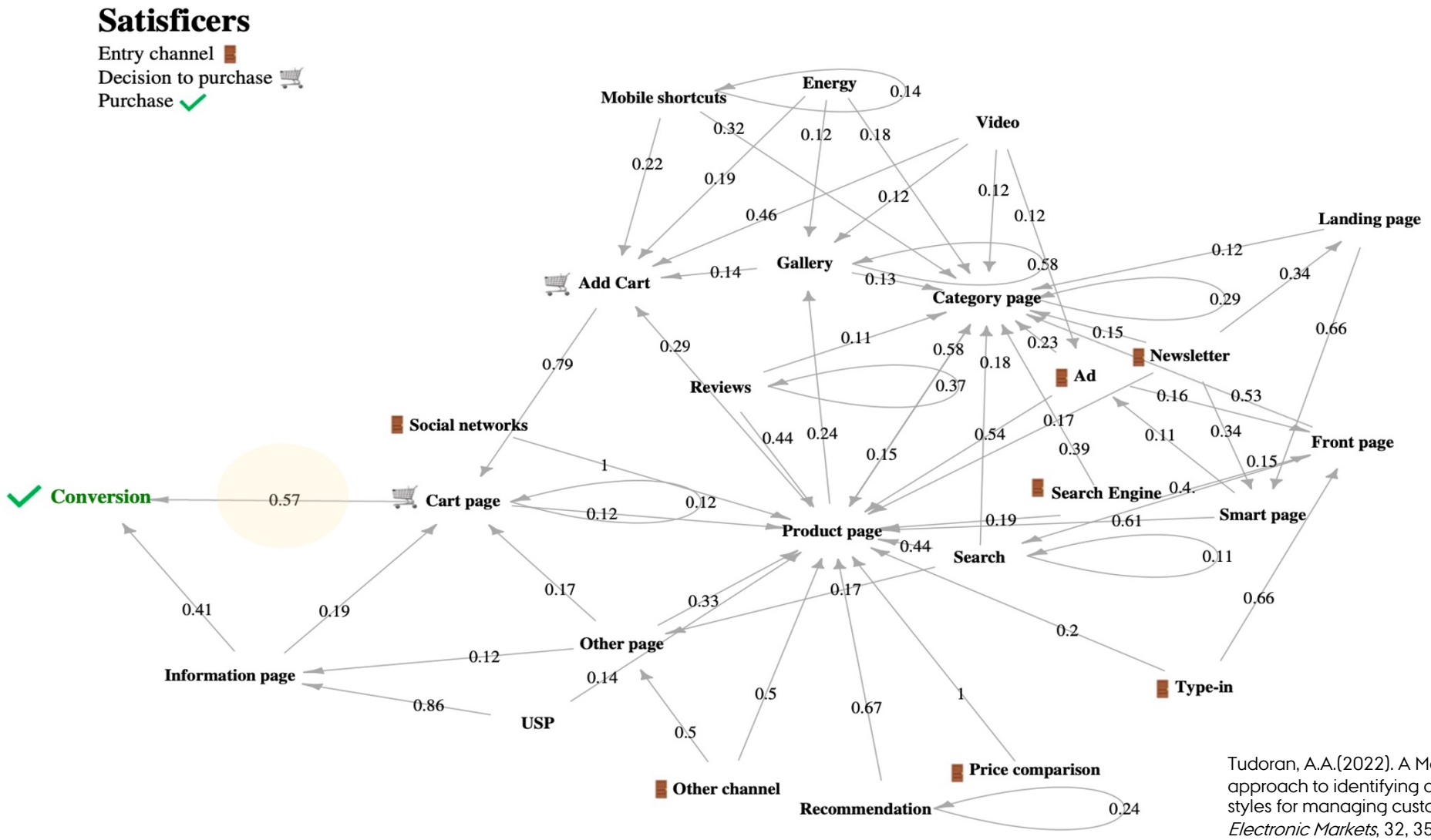
Aspect	Bayesian Networks	SEM
Definition	A graphical model that represents **probabilistic relationships** among variables using Bayesian inference	A statistical model that seeks to explain the s**tructural relationships** between multiple variables, both observed and unobserved (latent)
Primary Focus	**Causal discovery** and probabilistic prediction	Testing and estimating relationships based on a **hypothesized causal model**
Data Requirements	Can handle various types of data, but especially **categorical data**	Requires continuous data and **assumes multivariate normality**

COMPARATIVE OVERVIEW

Aspect	Bayesian Networks	SEM
Model Structure	**Directed Acyclic Graphs (DAGs)** where nodes represent variables and edges represent probabilistic dependencies	A combination of measurement and structural models, usually presented as **path diagrams** with observed and latent variables
Causality	Causality is guided by **both data and domain expertise**	Causality is defined by **the theoretical model.**

PREDICT CUSTOMER ONLINE BEHAVIOR

(a) The event transition probabilities of order 1 for satisficers



Tudoran, A.A.(2022). A Machine Learning approach to identifying decision-making styles for managing customer relationships, *Electronic Markets*, 32, 351-374.

PREDICT CUSTOMER PREFERENCES

BN are model-based algorithms for product recommendation (Collaborative filltering):

Item / User	I1	I2	I3	I4	I5
U1	Like	Dislike	Dislike		Like
U2	Dislike			Like	Dislike
U3		Like	Dislike		Like
Target user	Like	Dislike	Dislike	Like	?

A data set with individuals´s past behaviour (e.g. books visited, etc).

Person	X	Y	Z	W
Gloria	yes	no	yes	yes
Juan	no	yes	no	no
Amin	no	yes	yes	yes
Sam	yes	yes	yes	yes
Judy	yes	yes	yes	no
etc				

Suppose a new customer (Joe) is currently visiting some of these products.

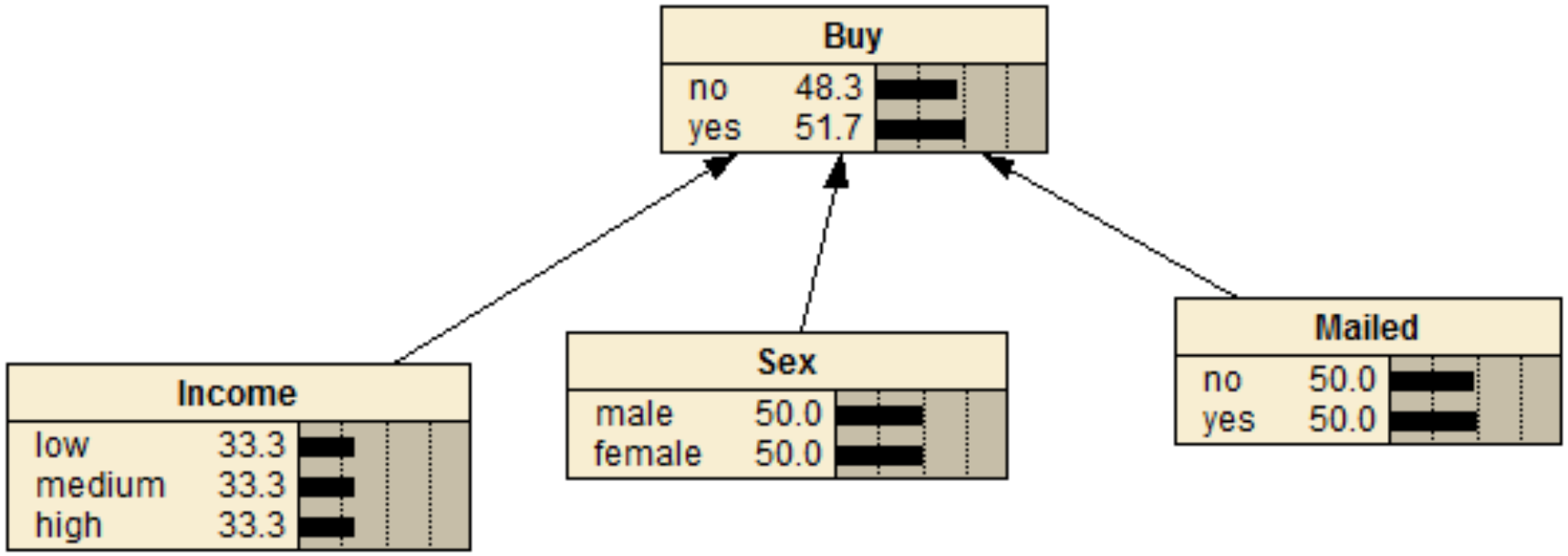
	X	Y	Z	W
Joe	yes	?	yes	?

Users	X	Y	Z	W
$user_1$	1	4	5	4
$user_2$	5	1	1	2
$user_3$	4	1	2	1
$user_4$	2	5	4	5
$user_5$	1	5	5	5
....				
	X	Y	Z	W
$user_a$	1	5	5	?

ANA ALINA TUDORAN
ASSOCIATE PROFESSOR

PREDICT CUSTOMER REACTIONS TO ADS

Targeted advertising based on probability to buy and the expected lift in profit :



PREDICT CUSTOMER CHOICE

1. Ask customers to choose from among product designs (something consumers do every day).
2. We decompose customer preferences into measurable attributes
3. Use a choice model to *infer* their preferences from the choices.
4. *Predict* whether they will buy alternative designs

Example from Chapman & Feit, *R for Marketing Research and Analytics, 2019*

EXAMPLE CHOICE QUESTION

Which of the following minivans would you buy?
Assume all three minivans are identical other than the features listed below.

	Option 1	Option 2	Option 3
	6 passengers 2 ft. cargo area gas engine $35,000	8 passengers 3 ft. cargo area hybrid engine $30,000	6 passengers 3 ft. cargo area gas engine $30,000
I prefer (check one):	☐	☐	☑

Fig. 13.1. An example choice-based conjoint survey question.

Respondents can answer up to 50 of these types of questions, which gives us lots of data from which to infer their preferences.

Example from Chapman & Feit, *R for Marketing Research and Analytics, 2019*

WHAT DRIVES DECISION?

Attribute Importance

- What Matters Most?
  - Passenger Number: High importance
  - Cargo Space: Medium importance
  - Engine Type: High importance
  - Price: Very high importance

Predicting the Market: By analyzing the utility scores, we can simulate potential market shares for each minivan configuration, helping us predict which features will drive future sales.

Q & A

EXPLAIN BEHAVIOR

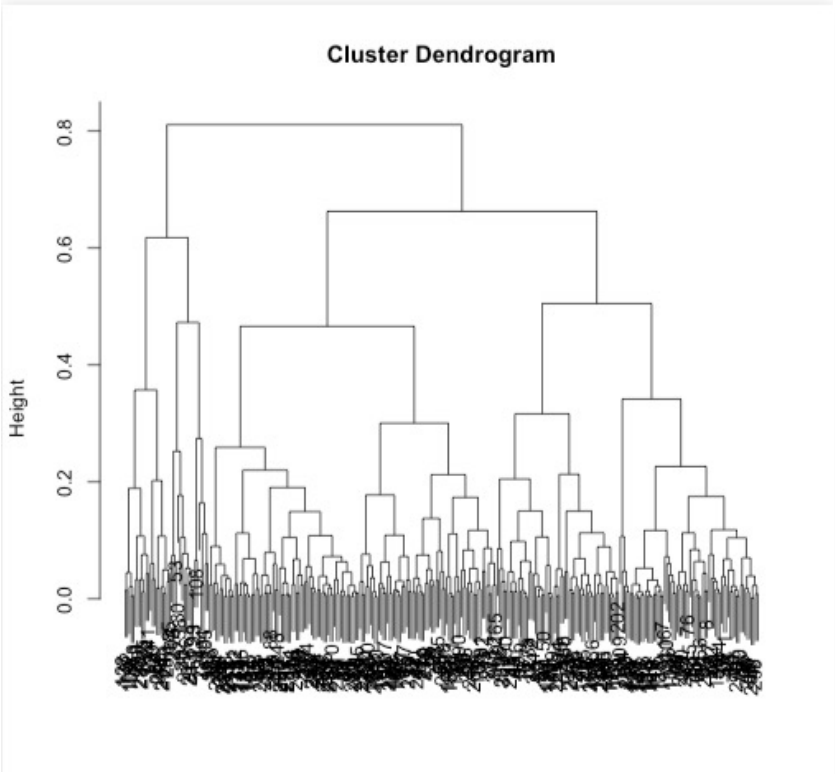
PREDICT CHOICES & BEHAVIOR

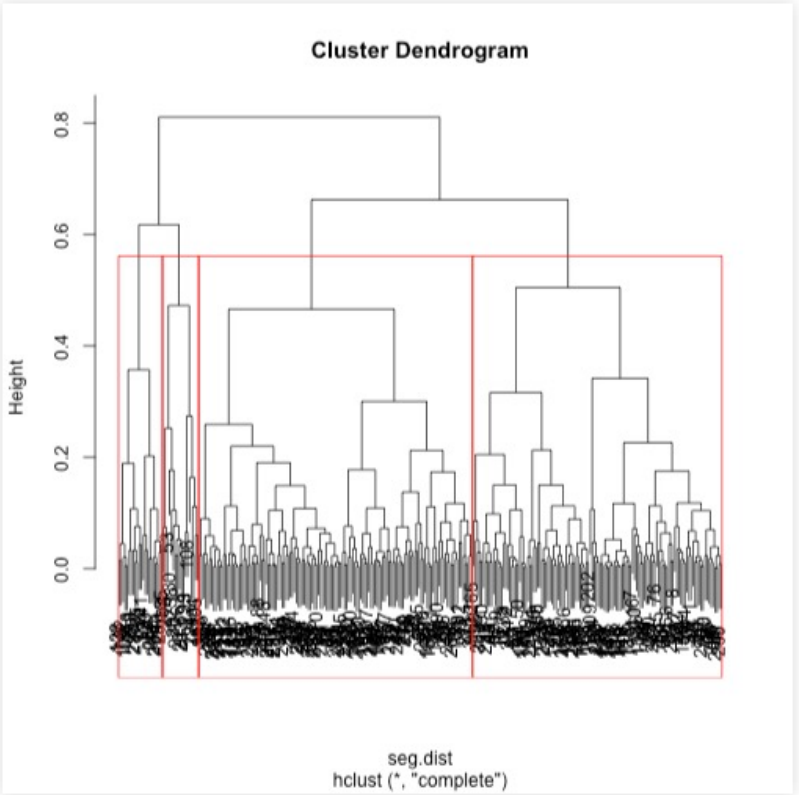
SEGMENT CUSTOMERS

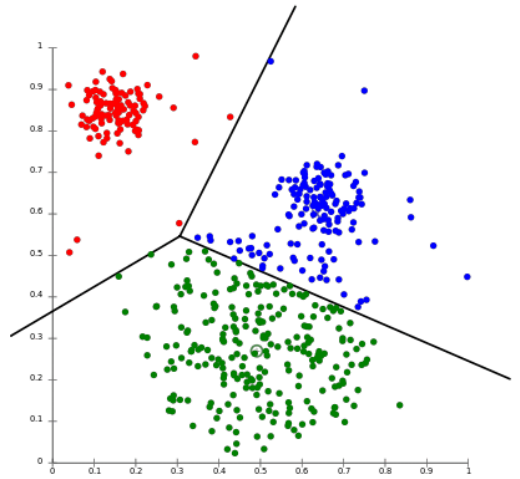
PREDICT RATINGS TO RECOMMEND PRODUCTS

CUSTOMER SEGMENTATION

Figures describe Hierarchical and K-means clustering







Algorithm 1 k-means algorithm

1: Specify the number k of clusters to assign.
2: Randomly initialize k centroids.
3: **repeat**
4: **expectation:** Assign each point to its closest centroid.
5: **maximization:** Compute the new centroid (mean) of each cluster.
6: **until** The centroid positions do not change.

COMPARATIVE OVERVIEW ON SEGMENTATION

Segmentation Aspect	K-Means Clustering	Hierarchical Clustering	Latent Class/Profile Analysis (LCA)	Gaussian Mixture Models	Bayesian Networks	Neural Networks
Methodology	Partitioning	Agglomerative and Divisive	Model-based Clustering	Expectation-Maximization	Probabilistic Graphical Models	Neural Networks
Nature of Technique	**Non-probabilistic** distance-based	**Non-probabilistic**, connectivity-based; requires to specify the # of clusters a-priori	**Probabilistic**, categorical (class) or continuous (profile)	**Probabilistic**, distribution-based	**Probabilistic**, causal inference-based	**Probabilistic** or **Non-probabilistic,** feature learning-based

EXPLAIN BEHAVIOR

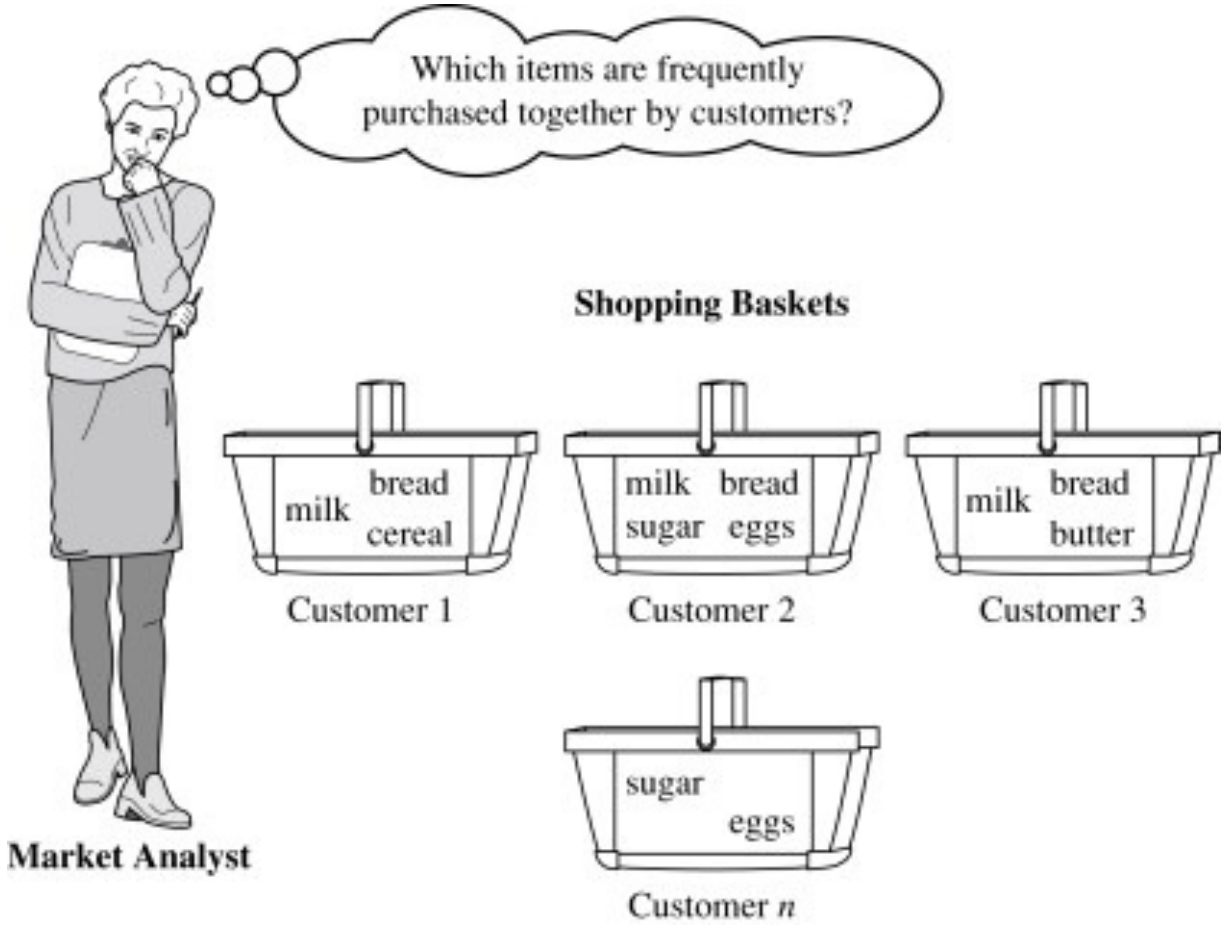
PREDICT CHOICES & BEHAVIOR

SEGMENT CUSTOMERS

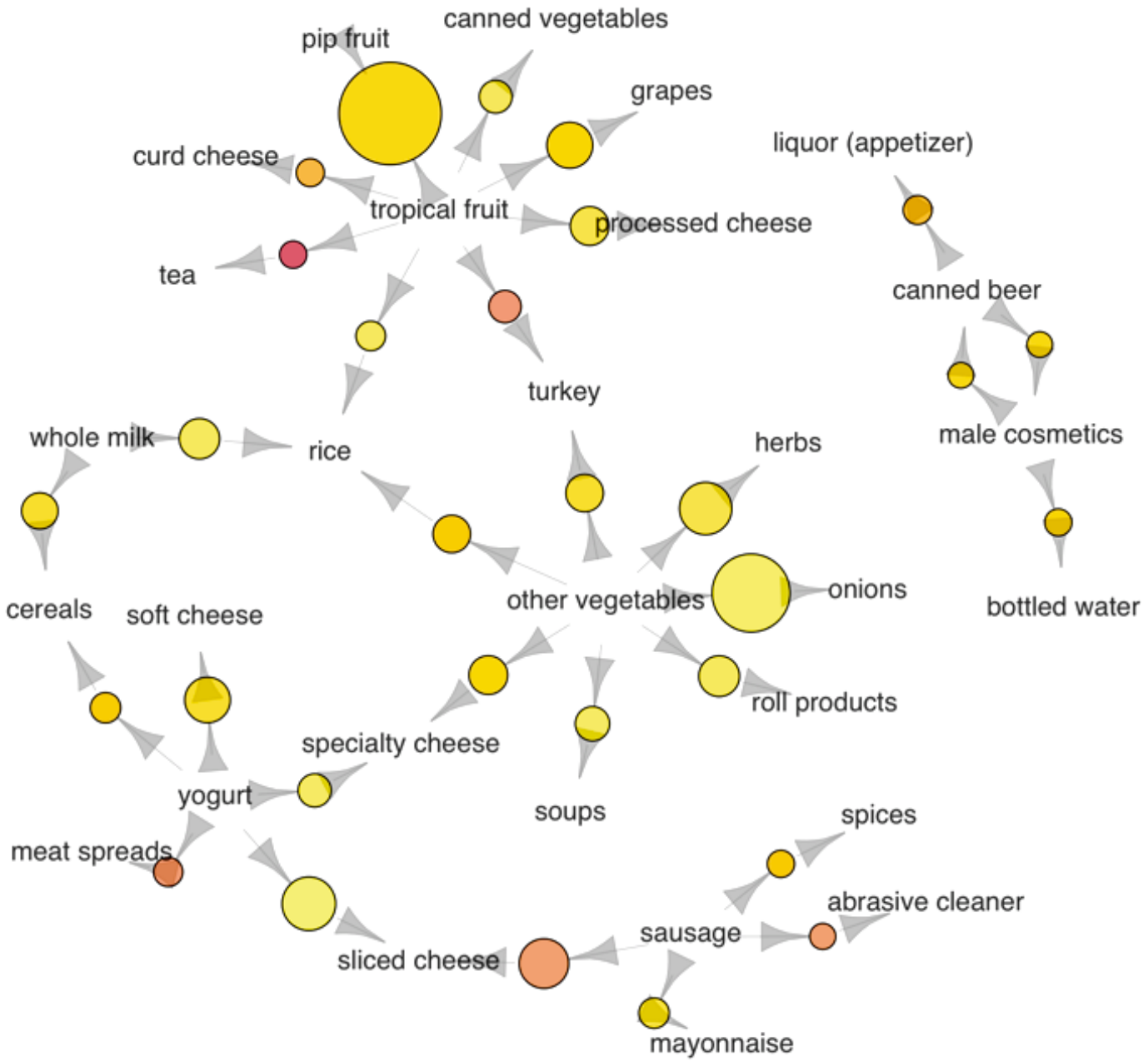
PREDICT RATINGS TO RECOMMEND PRODUCTS

ASSOCIATION RULE MINING (MARKET BASKET ANALYSIS)

A process that looks for relationships among objects that frequently appear together, such as the collection of items in a shopper's cart.



Han, J., Kamber, M., & Pei, J. (2012). 6-mining frequent patterns, associations, and correlations: Basic concepts and methods. In *Data mining* (pp. 243-278). Boston: Morgan Kaufmann.



https://www.kaggle.com/code/gauravchopracg/introduction-to-association-rule-mining

PRODUCT RECOMMENDATION (II)

- A user item rating matrix:

User \ Item	I1	I2	I3	I4	I5
U1	Like	Dislike	Dislike		Like
U2	Dislike			Like	Dislike
U3		Like	Dislike		Like
Target user	Like	Dislike	Dislike	Like	?

A descriptor, or a user rating

- The prediction task is to fill the target user´s empty elements of the matrix.
- Focus on the Cold Start Problem in Recommender Systems

OVERVIEW ON RECOMMENDER SYSTEMS

Aspect	Association Rule Mining	Memory-Based Recommendation	Model-Based Recommendation	Bayesian Networks for Collaborative Filtering
Methodology	Searches for interesting **correlations between products** in large datasets.	Uses simple memory-based rules to find **similarities between users or products**.	Employs **algorithms** to predict user preferences.	Utilizes **probabilistic graphical models** to predict user preferences.
Nature of Technique	Rule-based	User-item interactions and similarity calculations (e.g., **corr coefficient**).	Build a **training a predictive model**.	Build a training predictive probabilistic model

OVERVIEW ON RECOMMENDER SYSTEMS

Aspect	Association Rule Mining	Memory-Based Recommendation	Model-Based Recommendation	Bayesian Networks for Collaborative Filtering
Typical Data Used	**Transactional data**	User-item **ratings or user behavior data**	User-item **ratings or user behavior** data.	User-item ratings, user behavior data, **possibly additional contextual data.**
Advantages	Generates **easy-to-understand rules**	**Simple to understand** and implement; transparency in recommendation logic.	**Capture complex patterns;** more accurate with sufficient data.	Can handle uncertainty; incorporates domain knowledge; **provides insights into the reasoning behind recommendations.**

OVERVIEW ON RECOMMENDER SYSTEMS

Aspect	Association Rule Mining	Memory-Based Recommendation	Model-Based Recommendation	Bayesian Networks for Collaborative Filtering
Limitations	May produce **too many rules**; **doesn't capture the strength of preferences**.	May not handle sparse data well; **cold start problem.**	Requires careful **model selection and tuning; cold start problem.**	Requires **expert knowledge to set up the network**; potentially complex to interpret.
Common Uses	Cross-selling, product placement, inventory management.	E-commerce, small-scale recommendation engines.	E-commerce, streaming services, personalized content delivery.	Advanced recommendation systems, personalized marketing, decision support systems.

CONCLUSION

There is **not one method** in Customer Analytics **that stands above** the rest as the best. Each method is designed to address a specific need—be it understanding, predicting, segmenting, or recommending. Within each of these methods lies **a spectrum of algorithms**, each with its unique approach to solving the problem.

To understand the customers and make smart business decisions, one needs to know to implement several methods and more importantly, understand **when and why** to use each one.

REFERENCES

- Chapman, C., & Feit, E. M. (2019). *R for marketing research and analytics* (Vol. 67). New York, NY: Springer.
- Johnson, M. D., Gustafsson, A., Andreassen, T. W., Lervik, L., & Cha, J. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, 22(2), 217-245
- Tudoran, A.A.(2022). A Machine Learning approach to identifying decision-making styles for managing customer relationships, *Electronic Markets*, 32, 351-374.


UNIVERSITAS ARHUSIE
AARHUS
BSS

RECALL FOR NEXT LECTURE

Weeks 5(2) – 8(7). EXPLAINING CUSTOMER BEHAVIOR

FOR FACTOR ANALYSIS:

- Mehmetoglu M. and Mittner M. (2022). *Applied Statistics using R: A guide for the Social Sciences*, Sage. **Chapter 13.** Available on Brightspace.