Engineering the Knowledge Destination –

Applying Methods of Business Intelligence to the Management of Tourism Destinations
Agenda

• Introduction
  – Background and theoretical base
  – Research goal
• Knowledge destination framework architecture
• Excursion: Data Mining, Warehousing and Business Intelligence
• DMIS indicators for various types of data sources
  – Navigation-based
  – Transaction-based
  – Feedback-based
• DMIS Data Warehouse concept
  – Business processes
  – DMIS architecture
• DMIS Prototype
  – Feedback-based business process
• Literature
Introduction: The Knowledge Destination

• Background

  – Service societies = \( f \{\text{trends: societal, demoscopic, internationalization, new information and communication technologies (ICTS)}\} \) (Fitzsimmons & Fitzsimmons 2006) \( \Leftrightarrow \) Travel & Tourism (T&T)(Dwyer et al. 2007)

  – T&T

    • 5th world’s biggest economic sector (WTTC 2012)
      – $ 2 trillion in 2011 and 98 million employed people
      – double GDP if compared to automotive industries
      – Expected growth 2020: 4.2%

    • Information intensive \( \Leftrightarrow \) **T&T largest sector in e-Commerce**
      – 25.7% (i.e. € 65.2 Bn.) of EU online sales generated by T&T - in 2001: € 5 Bn. (Marcussen 2009)
Introduction: The Knowledge Destination

• Background
  – ICTS erode traditional firm boundaries and knowledge bases ⇔ behavioural and structural changes (market place → market space)
    • T&T ⇔ ICT shortcomings: collaborative offer generation, market research, knowledge creation, product innovation and strategic decision-making (E-Business Watch 2006)
    • Sustainable development and management of strategic unit in T&T - tourism destination - directly related to availability of knowledge needed to reconfigure networks of competencies and resources to assure competitiveness (Ritchie & Crouch 2003; Fesenmaier et al. 2004)
  – Knowledge-based view of the firm (Grant 1996)
    • Firm value is limited by amount of knowledge within it
Introduction: The Knowledge Destination

• Core competencies
  – Economic development of industries related to availability of knowledge needed to reconfigure ‘resources’ to remain competitive (Barney 1991: 101)
    • Only if knowledge resources are valuable to customers, scarce and difficult to imitate and substitute they establish competitive advantages
  – Entrepreneurial activity of combining and reconfiguring resources is based upon core competencies, which need to be renewed and validated through continuous learning and knowledge absorption processes
    • Dynamic capability ⇔ ‘firm’s ability to integrate, build and reconfigure internal and external competences to address changing environments’ (Teece et al. 1997: 516)
Introduction: The Knowledge Destination

• Capabilities of organizational learning
  – *Replication capability*
    • Effectively and efficiently multiplying established processes and operations \(\Leftrightarrow\) firm-internal knowledge transfer and codification processes
  – *Reconfiguration capability*
    • Continuously modifying existing resource configurations through acquisition and development of new core-competencies \(\Leftrightarrow\) determined by
      – absorbability of external knowledge (affected by ability to learn) and potential to deduce generalizable cause-effect relationships from existing knowledge applicable to a wider range of strategic options (Back et al. 2007)
      – firm’s proximity to the customer \(\Leftrightarrow\) relevance of customer-based knowledge (Tajeddini 2010)
Introduction: The Knowledge Destination

• **Sustainable knowledge sources**
  
  – emerge through generation, management and intelligent access of information relevant for resource stewardship and resource re-configuration (Back et al. 2007)

• **Competitiveness of tourism destinations**
  
  – depends on how information needs and organisational learning requirements of stakeholders can be satisfied (Shaw & Williams 2009)
  
  – **ICT-based infrastructures and services** play crucial role to increase stakeholders’ knowledge bases (i.e. ↓info asymmetries; Buhalis 2006)
  
  – Organisational learning and managerial effectiveness in tourism can be enhanced by applying **methods of business intelligence** (Pyo 2005)
Introduction: The Knowledge Destination

- **Business Intelligence** in the T&T domain
  - Makes individual knowledge about customers, products, processes, competitors or partners *meaningful and available*
  - Offers destination stakeholders reliable, up-to-date and strategically relevant knowledge (Wong et al., 2006, Fuchs & Höpken 2009)
  - Fosters large-scale intra & inter-firm knowledge exchange (Min et al. 2002)

- Large variety and quantity of **customer-based data** wide-spread in tourism destinations
  - However, huge amount remains *unused* (Höpken et al. 2011)
  - **Data bases** [transaction data, CRM data, survey data]
  - **Webservers** [navigation data, search data, UGC]
Objective

- Application of *business intelligence methods* (BI) to retrieve relevant and previously unidentified *knowledge* from *customer-based data*

- *Destination Management Information System (DMIS)* that supports value creation through enhanced decision making

  - Monitor fulfilment of strategic destination goals
  - Support for decision making to increase degree of strategic goal *fulfilment* (e.g. destination marketing, capacity planning, etc.)
Introduction: The Knowledge Destination

Knowledge application layer

Knowledge generation layer

Customer-oriented knowledge application
- Location-based services
- Community services

Supplier-oriented knowledge application
- Business intelligence applications (DMIS): Business performance management, forecasting, multi-channel management

Customer-based knowledge generation
- Tourists feedback
- Information traces
- Mobility behaviour

Supplier-based knowledge generation
- Products, processes, competitors and strategic partners

The Knowledge Destination Framework
(Höpken, Fuchs, Keil & Lexhagen 2011)
Intro summary: The Knowledge Destination

- Huge amounts of customer-based data in destinations unused
  - Transaction data
  - CRM data
  - Tracking data
  - Survey data
  - Navigation data
  - Search data [keywords]
  - UGC [ratings, blogs, e-reviews]
  - Marketing effectiveness
  - Quality of visitor experience
  - Economic performance

- Monitor fulfilment of strategic goals
  e.g. Are Vision 2020: satisfied guests, stays, contacts, brand awareness...

- Decision making support to increase degree of goal fulfilment
  e.g. media mix, reach, conversion, market basket, capacity forecast, cancellation...
Knowledge destination framework architecture

• Realizing the Knowledge Destination by
  – Application of methods of *business intelligence* (BI) to retrieve relevant and previously unidentified *knowledge* from *customer-based* data from *all destination value chain areas*

  • **Data Bases**: Transaction data, CRM data, Tracking data, Survey data
  • **WWW**: Navigation data, Search data, UGC [ratings, blogs, e-reviews]

  – **Destination Management Information Systems (DMIS Application)**
    • Data collection & information extraction
    • Data mining & knowledge generation
    • Knowledge application
Knowledge destination framework architecture

**Data sources**

<table>
<thead>
<tr>
<th>Explicit tourists’ feedback provided knowingly and intentionally</th>
<th>Implicit tourists’ information traces provided unknowingly and unintentionally</th>
</tr>
</thead>
<tbody>
<tr>
<td>- <em>Structured data</em>: e.g. online and offline guest surveys, ratings from web 2.0 applications, user profiles from web applications and online communities, etc.</td>
<td></td>
</tr>
<tr>
<td>- <em>Unstructured data</em>: free text from E-mails and web 2.0 applications (e.g. blogs, e-comments/reviews), rich content (e.g. YouTube.com), etc.</td>
<td>- <em>Navigation data</em>: search behaviour on web sites and online portals, community sites, etc.</td>
</tr>
<tr>
<td></td>
<td>- <em>Transaction data</em>: online requests, reservation and booking data, payment, etc.</td>
</tr>
<tr>
<td></td>
<td>- <em>Tracking data</em>: GPS/WLAN-based coverage of tourists’ spatial movements</td>
</tr>
<tr>
<td></td>
<td>- <em>Observation data</em>: gathered in a laboratory context or through market observation</td>
</tr>
</tbody>
</table>

Knowledge destination framework architecture

DMIS

Data mining & knowledge generation

Data warehouse

Data extraction (ETL)

Structured data

Unstructured data

Knowledge application layer

Knowledge generation layer

Data extraction

- **Structured data**: semantic, linguistic or constraint-based techniques
- **Html documents**: wrapper based on static patterns or (un)supervised learning methods
- **Free text**: text mining based on statistical language models or NLP

(Hoepken et al. 2011)
Knowledge destination framework architecture

Data warehouse
- Heterogeneous data stored in a **central** data warehouse in a **homogeneous structure**
- Individual **data sources mapped**
  1. into common **tourism ontology**
  2. into **central DW schema**

(Hoepken et al. 2011)
Knowledge destination framework architecture

Data mining & knowledge generation

- Recognition of **patterns and relationships** by data mining techniques
- Generation of knowledge in form of **validated models** (cluster models, classification models, association rules,...)

(Hoepken et al. 2011)
Knowledge destination framework architecture

Destination management information system (DMIS)

- Interactive visualization of
  - Validated data mining models
  - Underlying data by exploratory data analysis (EDA) and online analytical processing (OLAP)

(Hoepken et al. 2011)
Knowledge destination framework architecture

- **Data warehouse**
  - Data extraction (ETL)
  - Structured data
  - Unstructured data

- **Data mining & knowledge generation**

- **Knowledge generation layer**

- **Knowledge application layer**

- **DMIS**

- **Knowledge destination framework architecture**

  - **All-stakeholder encompassing** generation of knowledge relevant to DMO and suppliers through techniques of BI
  - Knowledge **interactively visualized** through DMIS accessible to all destination suppliers
Excursion: Data Mining

• **Data mining** is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition techniques as well as statistical and mathematical techniques (Gartner Group)
  
  – Weak data assumptions
    • Non-linearity
    • All distribution types
    • Uncomplete, redundant, noisy and dynamic
    • Huge in amount

• **Business intelligence**
  
  – Umbrella term for data warehousing, data mining and OLAP
Excursion: Data Warehousing

- **Data extraction, transfer and loading (ETL)**
  - Data identification, org and preparation
    - Huge amount of sources (i.e. internal enterprise level, external application level)
    - Focus on **customer and product characteristics** (e.g. quality, price, distribution source) versus transaction-oriented data-base system
  - Directly supports analyses by **database query** (e.g. MySQL; **Online Analytical Processing - OLAP**)
    - Rotation
    - Drilling Up/Down (Dicing)
    - Ranging

- „Flat“ data organisation for automated pattern recognition with Data Mining
Excursion: Data Mining – Why Data Mining?

• Data Mining = One of 10 technologies which will change the world (MIT Tech Review, 2001)

• Growing importance of data mining due to...
  – Explosive growth of data flows [B2B/C, Intra/Internet/WWW] and collection (e.g. shopping / payment, travelling, ...) ⇔ huge amount of info on markets, competitors and customers
  – WWW enable access to extensive customer online behaviour
  – Decreasing costs for storing data in data warehouses and data transfer in operative IT systems due to growth of computing power and storage capacity
  – Open-source SW for Artificial Intelligence applications (WEKA™; RapidMiner™)
  – Competitive pressure due to globalization ⇔ need for ad-hoc analyses from marketing managers

• Current situation
  – Possibilities of modern information technology for business intelligence and decision support are not yet used adequately ⇔ “we are drowning in information but starved for knowledge” (John Naisbitt)
Excursion: Data Mining in Tourism

• Cutting edge
  – Airlines, online platforms, travel intermediaries, resort management, hotel chains
  – Small-sized tourism enterprises (e.g. private accommodation and gastronomy sector) catching up due to small budgets for marketing and ICT

• Data mining applications in tourism
  – *Customer segmentation, product bundling and positioning*
    • Clustering with Artificial Neural (e.g. Kohonen) Networks, Genetic algorithms
  – *Customer Relationship Management* (i.e. discovery of attributes of those customers with above average propensity to buy and contribution to profitability)
    • Classifikation through Artificial Neural Networks, Decision Tree Analysis
  – *Demand forecast for Revenue and capacity management*
    • Time series analysis with Artificial Neural Networks
  – *Travel decision-making and buying behaviour, advertising effectiveness, guest loyalty*
    • Association rules with a-priori algorithms, genetic algorithms, KNN
Excursion: DM – Basic Types of Machine Learning

• **Supervised Learning**
  – Model is learnt through training data (e.g. classification, forecast)
  – Training data comprise result of classification, forecast → model is then ‘applied’ to new data with unknown target variable

• **Unsupervised Learning**
  – Model is learnt on the base of available data without any guidance (e.g. clustering, associations recognition)
  – Target variable is not known a-priori

• **Overspecification – the overfitting problem**
  – Learning each specific structure from training data on cost of generalization and stability of the model
Excursion: DM Main tasks within Business Intelligence
Excursion: Data Mining

1) Data Description

- **Describes simple patterns, trends and relationships** within data
  - Data base query (SQL), OLAP (online analytical processing) and visualisation
- **Exploratory data analysis (EDA)** -> big picture, however no too profound insights
  - explorative statistics (e.g. mean, sum, cor, std dev...)

Tourist arrivals Barcelona and Google search volume (Költringer 2009)
Excursion: Data Mining

2) Classification

- Estimation of a categorical target variable (classifier)
  - **Example**: determine, based on attributes, like search patterns, booked products, to which of three customer segments first time visitors belong to
  - Models are built using “complete” records, providing values for the target variable (classifications) as well as the predictors (input variables)

<table>
<thead>
<tr>
<th>row no.</th>
<th>Play</th>
<th>Outlook</th>
<th>Temperature</th>
<th>Humidity</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no</td>
<td>sunny</td>
<td>85</td>
<td>85</td>
<td>false</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>sunny</td>
<td>80</td>
<td>90</td>
<td>true</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>overcast</td>
<td>83</td>
<td>78</td>
<td>false</td>
</tr>
<tr>
<td>4</td>
<td>yes</td>
<td>rain</td>
<td>70</td>
<td>90</td>
<td>false</td>
</tr>
<tr>
<td>5</td>
<td>yes</td>
<td>rain</td>
<td>80</td>
<td>80</td>
<td>false</td>
</tr>
<tr>
<td>6</td>
<td>no</td>
<td>rain</td>
<td>85</td>
<td>70</td>
<td>true</td>
</tr>
<tr>
<td>7</td>
<td>yes</td>
<td>overcast</td>
<td>84</td>
<td>65</td>
<td>true</td>
</tr>
<tr>
<td>8</td>
<td>no</td>
<td>sunny</td>
<td>72</td>
<td>85</td>
<td>false</td>
</tr>
<tr>
<td>9</td>
<td>yes</td>
<td>sunny</td>
<td>99</td>
<td>70</td>
<td>false</td>
</tr>
<tr>
<td>10</td>
<td>yes</td>
<td>rain</td>
<td>75</td>
<td>80</td>
<td>false</td>
</tr>
<tr>
<td>11</td>
<td>yes</td>
<td>sunny</td>
<td>75</td>
<td>70</td>
<td>true</td>
</tr>
<tr>
<td>12</td>
<td>yes</td>
<td>overcast</td>
<td>72</td>
<td>90</td>
<td>true</td>
</tr>
<tr>
<td>13</td>
<td>yes</td>
<td>overcast</td>
<td>81</td>
<td>75</td>
<td>false</td>
</tr>
<tr>
<td>14</td>
<td>no</td>
<td>rain</td>
<td>71</td>
<td>80</td>
<td>true</td>
</tr>
</tbody>
</table>

Training data

Learned model
**Excursion: Data Mining**

2) Classification

- **Flexibility of supervised learning**: if linearity, normality and non-inter correlation assumption fails ⇔ **discriminant analysis/logistic regression** (rarely the case)
- Typical methods: k-nearest neighbour, decision tree, artificial neural networks, association rules

---

**Fig. 2. The neural network for “Will Revisit” loyalty measure.**


Excursion: Data Mining

3) Estimation

- Estimation of a numerical target variable (attributes or features)
  - Models are built using “complete” records, providing values for target variable and predictors (input variables)
  - For new observations, the value of the target variable is estimated based on values of the predictors
  - **Examples:** estimate tourists spending based on their income
  - Typical methods: point estimation, confidence interval estimation, regression, but also artificial neural networks (ANN)
3) Estimation

\[ y = 0.2077x + 227.478 \] (+ depreciation, interest and leasing costs)

\[ Y \text{ [Total Expenses]} = $359.429 \text{ [Fix Cost]} + 0.2077 \text{ [Var. Cost]} \times \text{sales} \]

\[
\text{Break-even sales} = \frac{\text{Fix Cost}}{1 - \left( \frac{\text{Variable Cost}}{\text{Sales}} \right)}
\]

Break Even Sales $453.653
Variable Cost (0.2077) $94.224
Fix Cost $359.429
Total Cost $453.653
Profit $-0-$
Excursion: Data Mining
4) Prediction

• Estimation of target variable in the future
  – Models are built using “complete” records, providing values for target variable and the predictors (input variables)
  – For future observations, the value of the target variable is estimated based on values of predictors
  – Examples: predict future airline ticket prices, tourism demand, etc.
  – Typical methods: statistical methods for estimation; data mining techniques like neural network, k-nearest neighbour, ...

![Graph showing passenger data and trend line](image)

<table>
<thead>
<tr>
<th>Year</th>
<th>Passengers (Tsd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>515</td>
</tr>
<tr>
<td>2003</td>
<td>534</td>
</tr>
<tr>
<td>2004</td>
<td>576</td>
</tr>
<tr>
<td>2005</td>
<td>610</td>
</tr>
<tr>
<td>2006</td>
<td>679</td>
</tr>
<tr>
<td>2007</td>
<td>754</td>
</tr>
<tr>
<td>2008</td>
<td>748</td>
</tr>
<tr>
<td>2009</td>
<td>752</td>
</tr>
<tr>
<td>2010</td>
<td>764</td>
</tr>
</tbody>
</table>

**Forecast:**

Passengers ($Y_{11}$) = $-70,821 + 35,633 \times 2011$ (Tsd.)

**Reality:**

831 (Tsd.)

**Forecast Error:**

$0.8\%$
Excursion: Data Mining
5) Clustering

- Grouping of records into classes of similar objects
  - Segmentation of records into homogeneous clusters with maximal similarity of objects within each cluster and minimal similarity of objects of different clusters
  - Records do not contain any target variable
- Unsupervised learning
- **Example:** Grouping of customers (i.e. market segments) based on similarity of attributes from variables which describe them (e.g. demographics, travel behaviour, etc.)
- Typical methods: k-means clustering, hierarchical clustering, Kohonen networks

![Self organizing Artificial Kohonen Network 40 x 40 and travel segments](image)
Excursion: Data Mining

6) Association

- Identification of attributes in variables which “go together”
  - Identification of rules for quantifying relationships between attributes
  - Association rule: if *antecedent* (e.g. reservation room type A) → then *consequent* (booking additional offer) ⇐ *market basket analysis*: which items are purchased together
  - Typical methods: a priori algorithm, Generalized Rule Induction (GRI)

<table>
<thead>
<tr>
<th>No.</th>
<th>Premises</th>
<th>Conclusion</th>
<th>Support</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a3 = range5</td>
<td>a4 = range5 [1.950]</td>
<td>0.140</td>
<td>0.700</td>
</tr>
<tr>
<td>2</td>
<td>a3 = range5</td>
<td>a1 = range5 [0.550]</td>
<td>0.100</td>
<td>0.714</td>
</tr>
<tr>
<td>3</td>
<td>a4 = range5</td>
<td>a3 = range5 [0.850]</td>
<td>0.140</td>
<td>0.724</td>
</tr>
<tr>
<td>4</td>
<td>a3 = range1</td>
<td>a4 = range1 [-∞ - 0]</td>
<td>0.180</td>
<td>0.730</td>
</tr>
<tr>
<td>5</td>
<td>a3 = range5</td>
<td>a4 = range5 [1.950]</td>
<td>0.100</td>
<td>0.750</td>
</tr>
<tr>
<td>6</td>
<td>a2 = range2</td>
<td>a1 = range3 [0.850]</td>
<td>0.107</td>
<td>0.792</td>
</tr>
<tr>
<td>7</td>
<td>a3 = range5</td>
<td>a4 = range1 [-∞ - 1]</td>
<td>0.113</td>
<td>0.773</td>
</tr>
<tr>
<td>8</td>
<td>a4 = range1</td>
<td>a3 = range1 [-∞ - 1]</td>
<td>0.180</td>
<td>0.794</td>
</tr>
<tr>
<td>9</td>
<td>a3 = range1</td>
<td>a4 = range1 [-∞ - 0]</td>
<td>0.113</td>
<td>0.810</td>
</tr>
<tr>
<td>10</td>
<td>a3 = range3</td>
<td>a4 = range3 [1.150]</td>
<td>0.167</td>
<td>0.882</td>
</tr>
<tr>
<td>11</td>
<td>a1 = range5</td>
<td>a3 = range5 [0.350]</td>
<td>0.100</td>
<td>0.882</td>
</tr>
<tr>
<td>12</td>
<td>a2 = range2</td>
<td>a4 = range3 [1.150]</td>
<td>0.107</td>
<td>1</td>
</tr>
</tbody>
</table>
Excursion: Data Mining

6) Association

- Parameters in market basket analysis
  - *Support*
    - Frequency of particular item combination in transaction data \( \Rightarrow \)
      Probability that two product items are bought together \( [P(A \cap B)] \)
  - *Confidence*
    - Strength of this relationship \( \Rightarrow \) Probability that product B will be bought preconditioned that product A was already bought \( [P(B \mid A)] \)
    - e.g. 50% \( \Rightarrow \) in each 2\textsuperscript{nd} transaction which contains A also B appears
    - \( P(A \cap B) / P(A)P(B) > 1 \)
      - Implies that product A and B show a positive relationship (association) to be bought together
### Excursion: Data Mining

6) Association

<table>
<thead>
<tr>
<th>Entrée: Support (%)</th>
<th>Appetizer/Soup: Confidence (%)</th>
<th>Starch: Confidence (%)</th>
<th>Drink: Confidence (%)</th>
<th>Dessert: Confidence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla chicken</td>
<td>Corn soup 31.98</td>
<td>Bread 84.01</td>
<td>Ice tea 21.00</td>
<td>Mango cheese 33.53</td>
</tr>
<tr>
<td>13.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice hamburger</td>
<td>Corn or onion soup 29.77</td>
<td>Bread 95.67</td>
<td>Ice tea 21.37</td>
<td>Mango cheese 31.81</td>
</tr>
<tr>
<td>10.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shrimp hamburger</td>
<td>Corn soup 27.81</td>
<td>Bread 76.69</td>
<td>Ice tea 25.00</td>
<td>Mango cheese 30.90</td>
</tr>
<tr>
<td>9.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spice chicken</td>
<td>Corn soup 7.46</td>
<td>Bread 80.58</td>
<td>Ice tea 17.27</td>
<td>Mango cheese 27.70</td>
</tr>
<tr>
<td>7.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steak scallop</td>
<td>Clam soup 24.54</td>
<td>Bread 87.18</td>
<td>Ice tea 19.41</td>
<td>Mango cheese 38.10</td>
</tr>
<tr>
<td>7.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicken fish</td>
<td>Onion soup 27.22</td>
<td>Bread 86.52</td>
<td>Hot coffee 17.98</td>
<td>Mango cheese 34.46</td>
</tr>
<tr>
<td>6.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spicy chicken</td>
<td>Onion soup 31.02</td>
<td>Bread 98.78</td>
<td>Hot coffee 18.78</td>
<td>Mango cheese 34.69</td>
</tr>
<tr>
<td>6.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shrimp steak</td>
<td>Onion soup 31.54</td>
<td>Bread 87.55</td>
<td>Orange juice 20.75</td>
<td>Mango cheese 37.76</td>
</tr>
<tr>
<td>5.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roast spaghetti</td>
<td>Corn soup 38.16</td>
<td>Bread 96.62</td>
<td>Orange juice 25.60</td>
<td>Mango cheese 32.37</td>
</tr>
<tr>
<td>5.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broadway sizzling</td>
<td>Corn soup 25.87</td>
<td>Bread 79.10</td>
<td>Orange juice 21.89</td>
<td>Mango cheese or pudding</td>
</tr>
<tr>
<td>5.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filet steak</td>
<td>Onion soup</td>
<td>Bread 79.10</td>
<td>Ice tea 21.89</td>
<td>Mango cheese</td>
</tr>
</tbody>
</table>
Excursion: DM – Web Data Mining

• The Web is largest publicly available data source in the world
  – 295 Exa-bites (2009)
  – 93% of Internet surfers use tourism websites when planning a trip
• Mining the web is useful (and fascinating😊) because
  – Amount of information is huge and coverage is wide and diverse
  – The web contains data of all types (structured data [5%], semi-structured web pages [html], unstructured [text, multimedia files, meta data, hyperlinks])
  – Information on the web is heterogeneous, linked and noisy (parts are not relevant or of low quality)
  – The web also provides services
  – The web is dynamic
  – The web is a virtual society

These characteristics represent challenges and opportunities for data mining in the web
Excursion: DM – Web Data Mining

• **Web data mining**
  – Aims to discover and extract useful info from the Web’s hyperlink structure, page content and usage data
  – Is not purely an application of DM but invented many new mining tasks and algorithms
  – Tourism online market research (based on Web data mining) still in its infancy

• **Web structure mining**
  – Discovers useful info from hyperlink structure in a specific domain or website (e.g. communities of users, categories of similar websites such as authorities [particularly content rich] or hubs [overview sites])

• **Web content mining**
  – Extracts useful info from web page contents (e.g. classification or clustering of web pages, discover customer sentiments through text mining)

• **Web usage mining**
  – Discovers web usage patterns/behaviour of web-users from web log file data (e.g. click-stream analysis, discovery of user preferences)
  – Hugh amount of data, data generation unobtrusive (unbiased knowledge) at particularly low costs (only memory costs)
Excursion: DM – Web Structure Mining

• Hyperlinks
  – Main positioning determinant through search engines
    • Page Rank: Probability for detection /traceability of web-sites

• Example: Island destination of Elba (2008) vs. WWW
  ![Diagram of web structure mining]

• Elba’s Page Rank Distribution ⇄ Linkphobia in Tourismus (Baggio 2009)
  \( \varnothing = 2.85 \)

- SCC: strongly connected component
- IN: connects SCC (NOT reachable by SCC)
- OUT: connected by SCC (cannot reach SCC)
- TENDRILS: not connected to SCC (reachable from IN & can reach OUT)
- TUBES: directed paths (IN → OUT)
- DCC: disconnected components
Excursion: DM – Web Content Mining

Business Intelligence through search word analysis (UGC)

www.visiteuropeancities.info (03-06, 5,500 search words for 186 EU cities)

Arrivals and Google search volume (Költringer 2009)

Wolk & Wöber, 2008

Profile: Guided Tours, attractions, events

Heidelberg, Lyon, Tallinn, Bergen, Kopenhagen, Amsterdam

Profile: Shopping, Attractions

Madird, Budapest, Prag, Nizza, Rotterdam

Profile: Guided Tours, Opera, Museum, Art

20 search words for 32 cities

Mid: No special focus

Wolk & Wöber. 2008
Excursion: DM – Web Usage Mining

- [www.visistare.se](http://www.visistare.se)
- Clustering \( \leftrightarrow \) Log files Aug 08 - Mar 09 (92,035 user sessions)
- X-Mean Clustering (2 \( \leq x \leq 30 \)) by category change

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td>8.7% (7,989)</td>
<td>3.1% (2,839)</td>
</tr>
<tr>
<td><strong>Duration</strong></td>
<td>12.6 min</td>
<td>12.2 min</td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td>Accom. (80%; a/s 6.4)</td>
<td>Eating (100%; a/s 8.4)</td>
</tr>
<tr>
<td></td>
<td>Program (29%; a/s 1.9)</td>
<td>To do (33%; a/s 1.6)</td>
</tr>
<tr>
<td></td>
<td>To do (26%; a/s 1.9)</td>
<td>To see (21%; a/s 0.8)</td>
</tr>
<tr>
<td><strong>Actions</strong></td>
<td>Tourism (11.2%)</td>
<td>Brochure (7.6%)</td>
</tr>
<tr>
<td></td>
<td>Congress (7.2%)</td>
<td></td>
</tr>
<tr>
<td><strong>Ext. Search</strong></td>
<td>Services (17.5%)</td>
<td>Eating (13.5%)</td>
</tr>
<tr>
<td></td>
<td>Accom. (10.5%)</td>
<td>Services (12.5%)</td>
</tr>
<tr>
<td></td>
<td>Activities (3.1%)</td>
<td>Activities (2.1%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Int. Search</strong></td>
<td>3.4 min</td>
<td>3.8 min</td>
</tr>
<tr>
<td></td>
<td>Accom. (22%)</td>
<td>Accom. (17.7%)</td>
</tr>
<tr>
<td></td>
<td>Skiing (5.5%)</td>
<td>Eating (12%)</td>
</tr>
<tr>
<td></td>
<td>Activities (5.1%)</td>
<td>Skiing (8%)</td>
</tr>
<tr>
<td><strong>Parametric</strong></td>
<td>5.1 min</td>
<td>4.3 min</td>
</tr>
<tr>
<td><strong>Used after</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>By</strong></td>
<td>21% (a/s 4.4)</td>
<td>19% (a/s 4.9)</td>
</tr>
</tbody>
</table>


**Excursion: DM – Web Usage Mining**

- [www.visistare.se](http://www.visistare.se)
- Clustering \( \leftrightarrow \) Log files Aug 08 - Mar 09 (92,035 user sessions)
- X-Mean Clustering (2 \( \leq x \leq 30 \)) by category change

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td>8.7% (7,989)</td>
<td>3.1% (2,839)</td>
</tr>
<tr>
<td><strong>Duration</strong></td>
<td>12.6 min</td>
<td>12.2 min</td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td>Accom. (80%; a/s 6.4)</td>
<td>Eating (100%; a/s 8.4)</td>
</tr>
<tr>
<td></td>
<td>Program (29%; a/s 1.9)</td>
<td>To do (33%; a/s 1.6)</td>
</tr>
<tr>
<td></td>
<td>To do (26%; a/s 1.9)</td>
<td>To see (21%; a/s 0.8)</td>
</tr>
<tr>
<td><strong>Actions</strong></td>
<td>Tourism (11.2%)</td>
<td>Brochure (7.6%)</td>
</tr>
<tr>
<td></td>
<td>Congress (7.2%)</td>
<td></td>
</tr>
<tr>
<td><strong>Ext. Search</strong></td>
<td>Services (17.5%)</td>
<td>Eating (13.5%)</td>
</tr>
<tr>
<td></td>
<td>Accom. (10.5%)</td>
<td>Services (12.5%)</td>
</tr>
<tr>
<td></td>
<td>Activities (3.1%)</td>
<td>Activities (2.1%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Int. Search</strong></td>
<td>3.4 min</td>
<td>3.8 min</td>
</tr>
<tr>
<td></td>
<td>Accom. (22%)</td>
<td>Accom. (17.7%)</td>
</tr>
<tr>
<td></td>
<td>Skiing (5.5%)</td>
<td>Eating (12%)</td>
</tr>
<tr>
<td></td>
<td>Activities (5.1%)</td>
<td>Skiing (8%)</td>
</tr>
<tr>
<td><strong>Parametric</strong></td>
<td>5.1 min</td>
<td>4.3 min</td>
</tr>
<tr>
<td><strong>Used after</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>By</strong></td>
<td>21% (a/s 4.4)</td>
<td>19% (a/s 4.9)</td>
</tr>
</tbody>
</table>
DMIS indicators

- Co-created knowledge relevant to destination success

<table>
<thead>
<tr>
<th>Explorative Analyses (OLAP)</th>
<th>Clustering</th>
<th>Association Rules/Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Navigation based indicators</strong></td>
<td><strong>Navigation based indicators</strong></td>
<td><strong>Navigation based indicators</strong></td>
</tr>
<tr>
<td>- Web navigation and channel use (e.g. page frequency, view time, path length, click-streams, etc.)</td>
<td>- Web usage-based clusters (keyword session or transaction-based)</td>
<td>- Sequential navigational patterns</td>
</tr>
<tr>
<td><strong>Transaction based indicators</strong></td>
<td><strong>Transaction based indicators</strong></td>
<td><strong>Transaction based indicators</strong></td>
</tr>
<tr>
<td>- Sales per accommodation type (price x overnight shifts)/per sending country/guest type (OLAP)</td>
<td>- Valuable tourist segments based on demographics and consumption behaviour</td>
<td>- Visitor forecast based on search volume</td>
</tr>
<tr>
<td>- Booking patterns per tourist activity / per tourism service</td>
<td>- Valuable tourist segments based on mobility behaviour in the destination</td>
<td>- Occupancy per accommodation type = f(marketing mix)</td>
</tr>
<tr>
<td>- Conversion rate per guest type / per sending country</td>
<td>- Conversion rate per data driven tourist segment</td>
<td>- Length of stay (average overnights/visitor) per sending country/guest type = f(marketing mix)</td>
</tr>
<tr>
<td><strong>Feedback-based indicators</strong></td>
<td><strong>Feedback-based indicators</strong></td>
<td><strong>Feedback-based indicators</strong></td>
</tr>
<tr>
<td>- CBBE metrics per sending county/per accommodation type</td>
<td>- Valuable tourist segments based on CBBE dimensions:</td>
<td>- UGC concerning Åre</td>
</tr>
<tr>
<td>- UGC (ratings)</td>
<td>- brand comprehension</td>
<td>- social network/interaction analysis (lead users)</td>
</tr>
<tr>
<td>- UGC concerning Åre</td>
<td>- functional dest. value areas</td>
<td></td>
</tr>
<tr>
<td>- positive/negative reviews --&gt; blog aggregator, sentiment detection</td>
<td>- emotional dest. value areas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- value for money assessment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- satisfaction &amp; loyalty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Social interaction clusters (based on UGC concerning Åre)</td>
<td></td>
</tr>
</tbody>
</table>

ICT in Tourism © copyright IFITT
DMIS indicators: navigation-based

Navigation Patterns

Forecast

Search-Term Analysis

### DMIS indicators: navigation-based

#### Navigation-based User Clusters

<table>
<thead>
<tr>
<th>Cluster Model</th>
<th>Attribute</th>
<th>cluster_0</th>
<th>Attribute</th>
<th>cluster_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 0:</td>
<td>accommodation</td>
<td>0.115</td>
<td>accommodation</td>
<td>1.512</td>
</tr>
<tr>
<td></td>
<td>whatson_start</td>
<td>0.121</td>
<td>accomm_appartment</td>
<td>0.922</td>
</tr>
<tr>
<td>Cluster 1:</td>
<td>booking</td>
<td>0.107</td>
<td>accomm_booking</td>
<td>0.621</td>
</tr>
<tr>
<td>Cluster 2:</td>
<td>directions</td>
<td>0.103</td>
<td>offers</td>
<td>0.583</td>
</tr>
<tr>
<td>Cluster 3:</td>
<td>activities_summer</td>
<td>0.091</td>
<td>booking</td>
<td>0.498</td>
</tr>
<tr>
<td>Cluster 4:</td>
<td>restaurant_bar</td>
<td>0.091</td>
<td>accomm_suite</td>
<td>0.389</td>
</tr>
<tr>
<td>Total number:</td>
<td>restaurant_bar</td>
<td>0.091</td>
<td>offers_private</td>
<td>0.259</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attribute</th>
<th>cluster_2</th>
<th>Attribute</th>
<th>cluster_3</th>
<th>Attribute</th>
<th>cluster_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>offers</td>
<td>1.308</td>
<td>pool_saunaworld</td>
<td>1.037</td>
<td>panorama_pictures</td>
<td>8.378</td>
</tr>
<tr>
<td>offers_privat</td>
<td>0.663</td>
<td>pool_sauna_pool</td>
<td>0.887</td>
<td>ownership_pan_pict</td>
<td>1.866</td>
</tr>
<tr>
<td>accommodation</td>
<td>0.319</td>
<td>spa</td>
<td>0.764</td>
<td>media</td>
<td>0.841</td>
</tr>
<tr>
<td>activities_summer</td>
<td>0.218</td>
<td>spa_treatments</td>
<td>0.516</td>
<td>accommodation</td>
<td>0.780</td>
</tr>
<tr>
<td>booking</td>
<td>0.186</td>
<td>pool_sauna_sauna</td>
<td>0.420</td>
<td>accomm_appartment</td>
<td>0.634</td>
</tr>
<tr>
<td>ownership</td>
<td>0.161</td>
<td>activities</td>
<td>0.295</td>
<td>pool_saunaworld</td>
<td>0.439</td>
</tr>
<tr>
<td>accomm_package</td>
<td>0.153</td>
<td>accommodation</td>
<td>0.245</td>
<td>offers</td>
<td>0.427</td>
</tr>
<tr>
<td>activities</td>
<td>0.118</td>
<td>offers</td>
<td>0.236</td>
<td>pool_sauna_pool</td>
<td>0.402</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Avg(session_length_minutes)</th>
<th>Avg.(clicks_per_session)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster_0</td>
<td>0.869</td>
<td>2.471</td>
</tr>
<tr>
<td>cluster_1</td>
<td>5.128</td>
<td>10.454</td>
</tr>
<tr>
<td>cluster_2</td>
<td>2.407</td>
<td>4.995</td>
</tr>
<tr>
<td>cluster_3</td>
<td>3.618</td>
<td>7.844</td>
</tr>
<tr>
<td>cluster_4</td>
<td>9.020</td>
<td>18.224</td>
</tr>
</tbody>
</table>
DMIS indicators: transaction-based

**Table 6. Bookings per customer segment**

<table>
<thead>
<tr>
<th>Segment</th>
<th>count(id)</th>
<th>Sum (visiting days)</th>
<th>Avg (visiting days)</th>
<th>Sum (persons)</th>
<th>Avg (persons)</th>
<th>Sum (rooms)</th>
<th>Avg (rooms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>individual_domestic</td>
<td>6887</td>
<td>16168</td>
<td>2.348</td>
<td>20394</td>
<td>2.961</td>
<td>7946</td>
<td>1.154</td>
</tr>
<tr>
<td>individual_foreign</td>
<td>2660</td>
<td>5238</td>
<td>1.969</td>
<td>8740</td>
<td>3.286</td>
<td>3178</td>
<td>1.195</td>
</tr>
<tr>
<td>company_domestic</td>
<td>1486</td>
<td>3224</td>
<td>2.196</td>
<td>8060</td>
<td>5.490</td>
<td>5104</td>
<td>3.477</td>
</tr>
<tr>
<td>TS guests</td>
<td>1450</td>
<td>3287</td>
<td>2.267</td>
<td>4994</td>
<td>3.444</td>
<td>1460</td>
<td>1.007</td>
</tr>
<tr>
<td>company_foreign</td>
<td>1433</td>
<td>3030</td>
<td>2.114</td>
<td>8927</td>
<td>6.230</td>
<td>4916</td>
<td>3.437</td>
</tr>
<tr>
<td>recurrent private</td>
<td>1051</td>
<td>2223</td>
<td>2.115</td>
<td>3530</td>
<td>3.359</td>
<td>1104</td>
<td>1.050</td>
</tr>
</tbody>
</table>

**Bookings / Origin ⇔ Conversion Rate**

<table>
<thead>
<tr>
<th>uri</th>
<th>country</th>
<th>overall_clicks</th>
<th>conversion rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>accommodation_booking</td>
<td>SWEDEN</td>
<td>1352</td>
<td>6.3%</td>
</tr>
<tr>
<td>accommodation_booking</td>
<td>NORWAY</td>
<td>719</td>
<td>9.2%</td>
</tr>
<tr>
<td>accommodation_booking</td>
<td>UNITED KINGDOM</td>
<td>513</td>
<td>24.3%</td>
</tr>
</tbody>
</table>

**Table 7. Customer segments by demographics and consumption behaviour**

<table>
<thead>
<tr>
<th>Cluster 0 (31343)</th>
<th>Cluster 1 (29774)</th>
<th>Cluster 2 (38883)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lodging period = week Customer type = private Booking source = phone + web Booking status = cancelled Product type = lodging Profile: private customer, booking a one week lodging via web and phone, often cancelling</td>
<td>Lodging period = short week / weekend Customer type = company / N/A Booking source = phone Product type = lodging + other Profile: short trip company customer, booking lodging + other products via phone</td>
<td>Lodging period = N/A Customer status = (non-registered) customer Customer type = private Booking source = phone Product type = ski rental Profile: not registered private customer, booking ski rental via phone</td>
</tr>
</tbody>
</table>
DMIS indicators: transaction-based

Cancellation by country

<table>
<thead>
<tr>
<th>Country</th>
<th>Cancellation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWE</td>
<td>13,4%</td>
</tr>
<tr>
<td>NOR</td>
<td>12,0%</td>
</tr>
<tr>
<td>FIN</td>
<td>19,3%</td>
</tr>
<tr>
<td>DNK</td>
<td>22,7%</td>
</tr>
<tr>
<td>RUS</td>
<td>31,1%</td>
</tr>
<tr>
<td>EST</td>
<td>28,7%</td>
</tr>
<tr>
<td>GBR</td>
<td>13,6%</td>
</tr>
<tr>
<td>NLD</td>
<td>21,6%</td>
</tr>
<tr>
<td>LVA</td>
<td>25,7%</td>
</tr>
<tr>
<td>LV</td>
<td>12,8%</td>
</tr>
<tr>
<td>DEU</td>
<td>11,9%</td>
</tr>
<tr>
<td>EE</td>
<td>13,4%</td>
</tr>
</tbody>
</table>

Cancellation trend

<table>
<thead>
<tr>
<th>Year</th>
<th>Cancellation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>22,6%</td>
</tr>
<tr>
<td>2005</td>
<td>19,5%</td>
</tr>
<tr>
<td>2006</td>
<td>19,6%</td>
</tr>
<tr>
<td>2007</td>
<td>21,1%</td>
</tr>
<tr>
<td>2008</td>
<td>15,6%</td>
</tr>
<tr>
<td>2009</td>
<td>12,7%</td>
</tr>
</tbody>
</table>

5% of customers cancellation rate 60% (days to arrival > 41 days, private, stay 7-9 days, booking logis) ⇔ Target Marketing

DMIS indicators: feedback-based

Tourist-Destination Brand relationship approach

- **Brand contacts** (input from destination)
  - **Destination awareness**
    - I see a lot of ads about Åre
    - I often read about Åre in newspapers and magazines
    - I often find information about Åre on the Internet
    - I have heard about Åre from friends and relatives
    - Åre is a famous site for international winter sports competitions

- **Destination attributes**
  - Skiing
    - Snow reliability, number and variety of ski slopes, overall quality of alpine skiing, overall quality of skiing experience, ski lifts
  - Activities
    - Non-ski winter activities, after ski, shopping
  - Service
    - Staff in accommodation facilities, quality of food and beverages, friendliness and professionalism of employees
  - Atmosphere
    - Easy skiing for children, family-friendly, clean and tidy, safe and secure
  - Value for money
    - For the ski experience, overall compared to other skiing destinations, reasonable prices
  - Other tourists
    - I liked the behavior of other tourists, it was easy to interact and communicate with other tourists

- **Value-in-Use**
  - Satisfaction
  - Enjoyment
  - Thrill
  - Fun and excitement
  - Joy of achievement

- **Loyalty**
  - I would still come to Åre even if it is more expensive than other ski resorts
  - I will come back to Åre in winter within 2 years
  - I consider Åre to be my first choice of a ski resort
  - I will encourage friends and relatives to visit Åre in winter

**Tangible and intangible destination attributes**
- **Skiing**
- **Activities**
- **Service**
- **Atmosphere**
- **Value for money**
- **Other tourists**

ICT in Tourism © copyright IFITT
DMIS indicators: feedback-based

DMIS indicators: feedback-based

• Goal
  – Real time customer feedback in DMIS
  – Real time statistics on guest mix and cold beds in DMIS
  – Large pool of guest contacts for guest survey (after visit)

• e-CRST Feedback panel at destination to collect content
  – Customer profile data
    • Arrival & departure date, travel group
    • Gender & age, country of residence
  – Information about the visit
    • Main visitation motive
    • Type of accommodation, accommodation area, accommodation provider
    • Type(s) of activity at destination
  – Ad-hoc customer feedback
    • Positive/Negative experiences (unstructured)
    • Overall satisfaction level (structured)
  – E-mail addresses and permission to participate in B (i.e. name for entering lottery)
Survey registration

Particulars of your stay

Date of arrival *

month / day / year

Date of departure *

month / day / year

Number of adults in your travel group

1

Number of children in your travel group

0

How would you describe the motive for your visit to Åre?

- Main reason for the trip
- An important stop on a longer trip
- Quick stop on a longer trip

Which of the following activities have attracted you to Åre this summer?

- Aviation
- Bicycling
- Climbing
- Events
- Fishing
- Golf
- Guided tours
- Hiking
- Horse riding
- Mountain biking
- Riding the cable car
- Shopping
- Sightseeing
- Spa and pool
- Sunbathing and swimming
- Water sports
Survey registration

Type of accommodation *
- Hotel
- Camping
- Rental cottage
- Own cottage
- Rental apartment
- Own apartment
- Accommodation owned by relatives or friends
- Other

Accommodation area
- Are village
- Björnen
- Tegelfjäll
- Duved
- Other

Name of accommodation / service provider

Gender *
- Male
- Female

Year of birth *

Country of residence *
- Denmark
- Estonia
- Finland
- Germany
- Latvia
- Lithuania
- Netherlands
- Norway
- Poland
- Russia
- Sweden
- UK
- Other country
Survey registration

Would you like to tell us about your experience during your recent stay in Åre?

Positive experience

[Text box for positive experience]

Negative experience

[Text box for negative experience]

To what extent are you satisfied with your recent stay in Åre?

1 2 3 4 5

Not satisfied at all 🖖 🖖 🖖 🖖 🖖 Highly satisfied

Your e-mail address *

[Input field for e-mail address]

confirm e-mail address

The survey is fully anonymous and your name will be stored separately from your responses. However, if you are the lucky winner of one of the lottery prizes, we kindly ask you to provide your full name.

Please enter full name

First name

Surname

Your registration was ...

[Submit button]
DMIS indicators: Summary

Destination Performance Indicators by DMIS-Åre
The Åre Data Warehouse: Business Processes and OLAP Analyses

INDICATOR TYPES

1. Economic Performance
   - 1.1. Bookings
   - 1.2. Overnights
   - 1.3. Prices
   - 1.4. Turnover

2. Customer Behaviour
   - 2.1. Website Navigation & Search
   - 2.2 Information & Booking Behaviour
   - 2.3 Consumption Behaviour
   - 2.4 Customer Profiling

3. Customer Perception & Experience
   - 3.1 Brand Awareness
   - 3.2 Perception of Destination Value Chain Areas
   - 3.3. Value for Money & Customer Satisfaction

Basic business events/transactions generating data

DMIS Structure
The Åre Data Warehouse: Business Processes and OLAP Analyses

Business Process: Booking

Fact-Dimension Table: Booking

OLAP analysis on any combination of dimensions and restrictions
The Åre Data Warehouse: Business Processes and OLAP Analyses

Business Process: Feedback

Fact-Dimension Table: Feedback

Feedback items can be changed at run time $\Leftrightarrow$ new questionnaires/items added any time

OLAP analysis on any combination of dimensions and restrictions

Guest satisfaction over time

Brand visibility of Åre among age groups

ICT in Tourism © copyright IFITT
The Åre Data Warehouse: Business Processes and OLAP Analyses

• OLAP Analyses across business processes
  – Info Request ↔ Booking
    • Which info channels generate most bookings for the sub-areas of Åre?
  – Booking ↔ Consumption
    • Which additional products are bought by guests booked in commercial vs. cold beds? (e.g. at certain periods, such as during events/festivals at Åre)?
  – Booking ↔ Feedback
    • Which brand dimensions are most relevant to attract bookings/overnights?
  – Consumption ↔ Capacity
    • What is the share of occupancy for cold beds?
  – Consumption ↔ Feedback
    • Do loyal guests spend more?
# The Åre Data Warehouse: Business Processes and OLAP Analyses

## Business Processes, their Granularity and related Dimensions

<table>
<thead>
<tr>
<th>Business Process</th>
<th>Fact Table</th>
<th>Dimension</th>
<th>Time</th>
<th>Date</th>
<th>Customer</th>
<th>Customer Use Profile</th>
<th>Customer Demo Profile</th>
<th>Product</th>
<th>Vendor</th>
<th>Supplier</th>
<th>Channel</th>
<th>Location</th>
<th>Feedback</th>
<th>Var Feedback</th>
<th>URI</th>
<th>Event</th>
<th>Session</th>
<th>Referral</th>
<th>Survey</th>
<th>Marketing</th>
<th>Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info Request</td>
<td>info request</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Navigation</td>
<td>click fact table</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>session fact table</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Booking</td>
<td>booking</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stay</td>
<td>stay</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption</td>
<td>consumption</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location Tracking</td>
<td>location poi</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>feedback</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fact Table

Feedback Fact Table
- Date Dimension
- Channel
- Customer Usage Profile
- Customer
- Customer Demographic Profile
- Location
- VarFeed

Feedback Fact
- QuestionnaireNo: INT (DD)
- FeedbackValue: SMALLINT

Time Dimension
- Product
- Supplier
- Feedback
- Survey
- Vendor
- Departure Week
- Arrival Week

ICT in Tourism
© copyright IFITT
DMIS Architecture

Data sources
- Text files
- Databases
- Html files

BI server
- Rapid Analytics
- Data warehouse
- Meta-data repository

DMIS Cockpit (OLAP, reporting)

RapidMiner (data mining)

Data
- All data sources
- And data types

Data Bases

ETL

Web-Service

DMCockpit
- HTML/CMS

RAPID ANALYTICS

MySQL

Are Data Warehouse

1&1
- Dedicated Server
  - 32 Cores
  - 64 GB Main memory
  - 2400 GB HDD

ICT in Tourism © copyright IFITT

International Federation for IT and Travel & Tourism
Live presentation DMIS Cockpit

http://www.dmis-are.com/beta
Live presentation DMIS Cockpit

<table>
<thead>
<tr>
<th>indicator</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>total visits (sessions)</td>
<td></td>
</tr>
<tr>
<td>unique visitors</td>
<td></td>
</tr>
<tr>
<td>total page views</td>
<td></td>
</tr>
<tr>
<td>average pages per visit</td>
<td></td>
</tr>
<tr>
<td>average visit time in seconds (session length)</td>
<td></td>
</tr>
</tbody>
</table>

**DMIS-Åre dashboard: Web Navigation process, log-file data**
Live presentation DMIS Cockpit

DMIS-Åre OLAP: Web Navigation process, log-file data
Live presentation DMIS Cockpit

**Booking**
This business process is still under construction. Stay tuned!
Please press button to see bookings in relation to sessions

DMIS-Åre across-business analysis (process: Booking)
Live presentation DMIS Cockpit

DMIS-Åre dashboard: feedback process, winter survey data
Live presentation DMIS Cockpit

DMIS-Åre OLAP: feedback process, summer survey data
Live presentation DMIS Cockpit

DMIS-Åre Dashboard: UGC-based satisfaction score by hotels and product areas (process: Feedback)
Live presentation DMIS Cockpit

DMIS-Åre OLAP – UGC Analyser (process: Feedback)
Outlook

• In its present version DMIS considers customer-based data only
• But it is planned in future research to integrate also supplier-based data sources from the entire digital eco-system of the destination Åre (e.g. information on products, processes and collaboration partners extracted from sources [i.e. web-sites] in the form of product profiles and availability information [e.g. booking engines]
• Valuable knowledge about suppliers’ service potential (property status), the complementarity of destination offers (on the base of market basket analyses), and their evaluation through tourists’ feedback will be gained
Outlook

• Moreover, in the near future, DMIS cockpit will also provide data mining processes, like classification, clustering, or prediction executed by the RapidMiner® data mining software

• A final future research goal comprises the application of real-time Business Intelligence to gain real-time knowledge on tourists’ on-site behavior as a valuable knowledge input for intelligent ubiquitous e-CRM applications in tourism destination


Used Literature


Used Literature


Used Literature


Used Literature


