

Data Mining im e-commerce am Beispiel der DB AG

Katja Steuernagel

GIP AG

SAS Institute GmbH

Universität Kaiserslautern

<http://www.katja-steuernagel.de>

Das Problem

- **Webseite der Deutschen Bahn:**
400.000 Besucher pro Tag
- **Online – Fahrkartenverkäufe:**
3.000 pro Tag (**1%** der Besucher)
- **Aber : ca. jeder 10. Besucher von „Surf&Rail“ kauft auch eine solche Fahrkarte !**

Gegenstand der Arbeit

- Welche Kunden kaufen die Fahrkarten im Internet?
- Ist das Verhalten der einzelnen Kundensegmente vorhersagbar?
- Welche Qualität hat diese Vorhersage?
- Was kann man über die Kommunikation zwischen Kunde und Verkäufer aussagen?

„Surf&Rail“



[www.bahn.de](#) | [Fahrpläne](#) | [Buchen & Kaufen](#) | [Angebote](#) | [Service](#) | [International Guests](#) | [Konzern](#) | [Presse](#)

[FahrkartenShop](#) | [BahnCard](#) | [Touristik](#) | [BahnShop 1435](#)

surf&rail-Ticket - Verbindungsanfrage

[Kontakt](#) | [Hilfe & Hinweise](#)

surf&rail: Das Internet-Ticket für ausgewählte Verbindungen.
Einfach buchen und selbst ausdrucken.

Unsere Angebotskonditionen:
- Pauschalpreis für Hin- und Rückfahrt 2. Klasse inkl. Sitzplatzreservierung
- Nur gültig in den gebuchten Zügen und auf den gebuchten Plätzen
- Kein Umtausch, keine Erstattung
- Personengebunden, nicht übertragbar

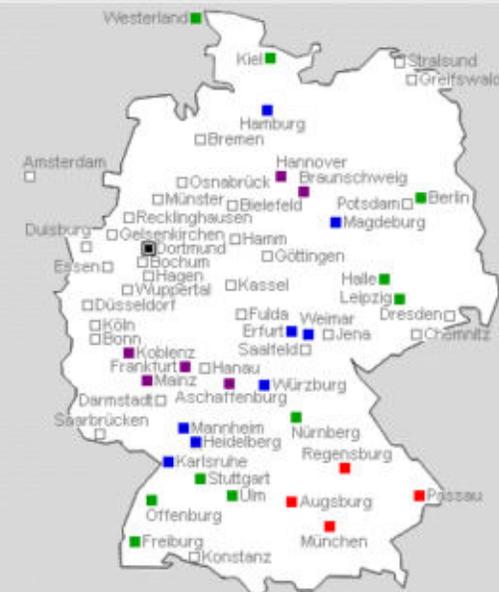
von:	<input type="text" value="Dortmund Hbf"/>
nach:	<input type="text" value="Aschaffenburg"/>
Hinfahrt Datum:	<input type="text"/> Buchung spätestens 7 Tage vor Abreise
Uhrzeit:	<input type="text" value="12:00"/> <input checked="" type="radio"/> Abfahrt <input type="radio"/> Ankunft
Rückfahrt Datum:	<input type="text"/> Frühestens Sonntag nach der Hinfahrt
Uhrzeit:	<input type="text" value="12:00"/> <input checked="" type="radio"/> Abfahrt <input type="radio"/> Ankunft
Personen:	<input type="text" value="1"/> Gemeinsame Platzreservierungen für Mitreisende ohne surf&rail-Buchung können derzeit nicht durchgeführt werden.

Verbindungen suchen

Preis pro Person mit/ohne BahnCard:

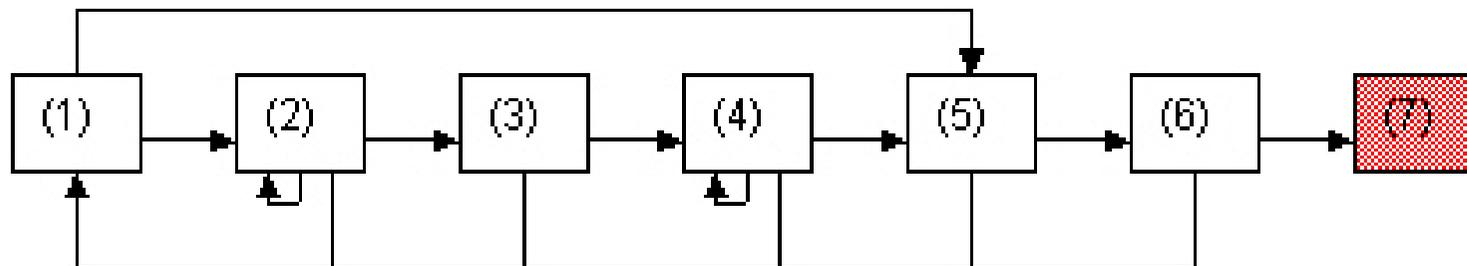
Durch Klick auf eine Stadt wählen Sie Ihren Abfahrtsbahnhof - und wir zeigen Ihnen die möglichen Ziele mit zugehöriger Preisstufe.

■ = 94/99 EUR ■ = 72/77 EUR □ = Abfahrt
■ = 50/55 EUR ■ = 35/40 EUR ○ = kein Zielbahnhof



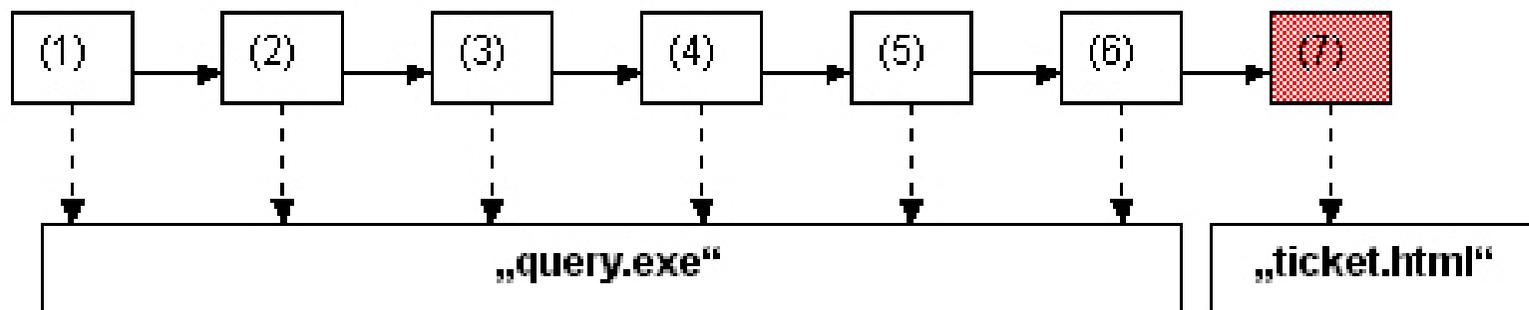
Schwierigkeiten:

- Reserviere 1518 (Wg.Art 03,00,00; NR) am 25.09.2001 von 8000284 nach 8010406
- **Schlechtes Webdesign**



Schwierigkeiten:

- Reserviere 1518 (Wg.Art 03,00,00; NR) am 25.09.2001 von 8000284 nach 8010406
- **Schlechtes Webdesign**



Schwierigkeiten:

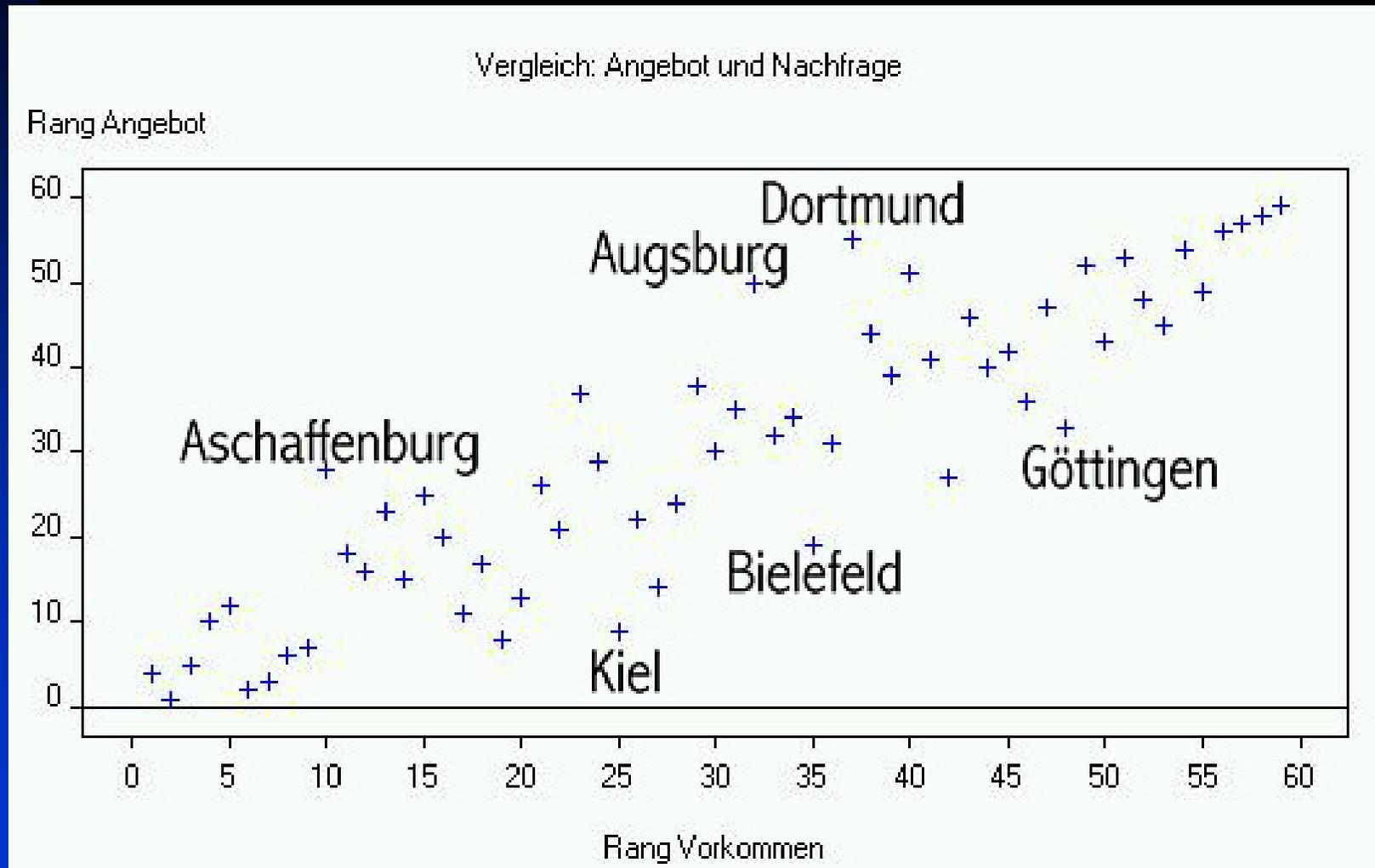
- Reserviere 1518 (Wg.Art 03,00,00; NR) am 25.09.2001 von 8000284 nach 8010406
- Schlechtes Webdesign

Schwierigkeiten:

- Reserviere 1518 (Wg.Art 03,00,00; NR) am 25.09.2001 von 8000284 nach 8010406
- Schlechtes Webdesign
- **Sonstige Probleme**



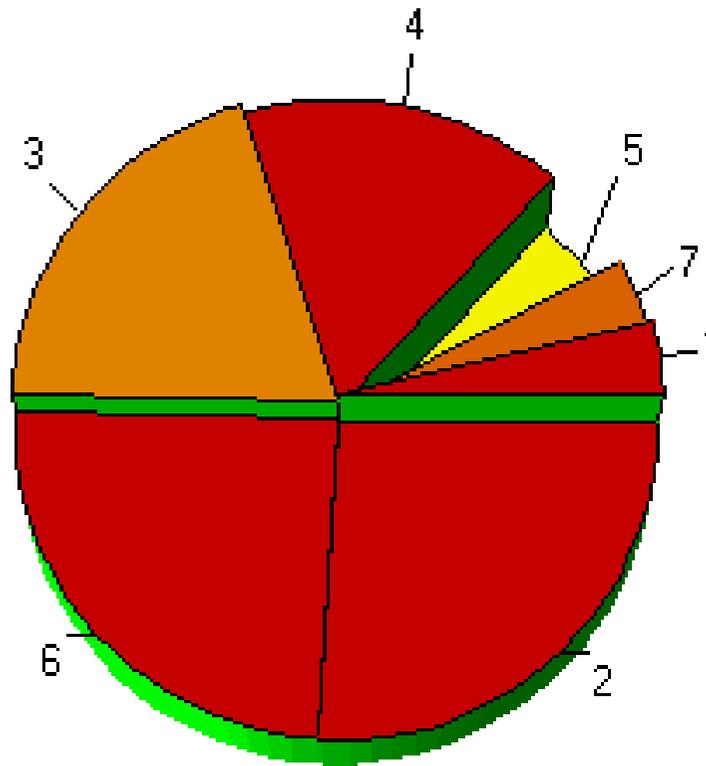
Vergleich Angebot & Nachfrage

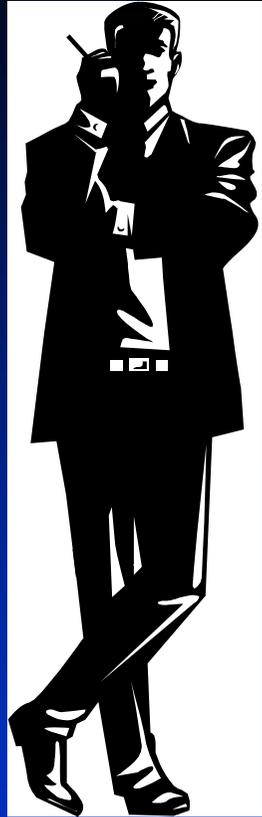


Besucherpfade

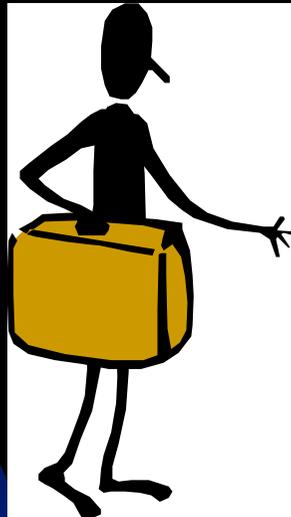
Java-Applet: [Besucherpfade](#)

Käufertypen





ICE,
Handyzone

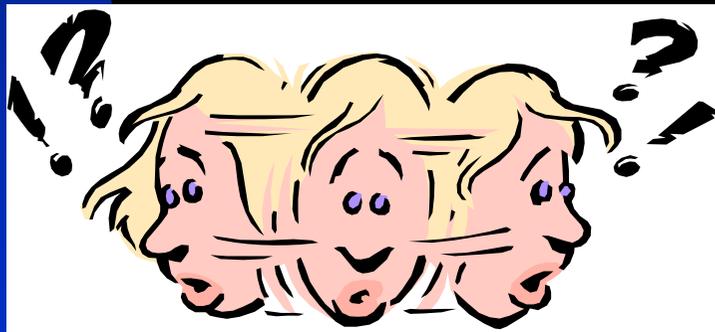


Single,
Wochenende



Single längere Reise

Käufertypen



???



Kurze Reise in der Gruppe

Nächste Schritte

- ***Vergleich S&R ↔ andere Fahrkarten:***
Gibt es Unterschiede in den Profilen und Verhaltensweisen der Benutzer?
- ***Sequenzanalyse:***
Wodurch unterscheiden sich Käufer und Nichtkäufer?
- ***Clustering:***
Sind erste Ergebnisse verbesserbar?
(eigene Algorithmen)

Dazu demnächst mehr ...

<http://www.katja-steuernagel.de>

