

TravelTide Kundensegmentierung

Datengetriebene Personalisierung für Kundenbindung & Umsatzwachstum

Projektbericht – November 2025

Claudia Tagbo | Data Science & Strategy

Inhaltsverzeichnis

Zusammenfassung	2
1 Methodik & Feature-Engineering	2
2 Segmentierungsansätze	3
2.1 Regelbasierte Segmentierung	3
2.2 ML-basierte Segmentierung (K-Means)	3
3 Ergebnisse & Perk-Performance	4
3.1 Perk-Verteilung	4
3.2 Wirksamkeit der Perks	4
3.3 Demografische Überschneidungen	5
4 Validierung & Geschäftsimpact	5
5 Empfehlungen & Roadmap	6
5.1 Phase 1: Sofortmaßnahmen (Wochen 1-4)	6
5.2 Phase 2: Optimierung (Monate 2-3)	6
5.3 Phase 3: Skalierung (Monate 4-6)	6
Fazit	6
Anhang	6

Zusammenfassung

TravelTide verfügt über eine große digitale Basis: **1,02 Millionen registrierte Nutzer**, die **5,4 Millionen Sitzungen**, **1,9 Millionen Flugbuchungen** und **1,9 Millionen Hotelreservierungen** generieren. Dennoch fehlt ein systematischer Ansatz zur Kundensegmentierung. Dieses Projekt schafft ein datenbasiertes Segmentierungsframework, das aus Verhaltensdaten handlungsorientierte Personas ableitet und personalisierte Vorteile („Perks“) zuweist.

Ziel: Entwicklung eines hybriden Segmentierungssystems – kombiniert regelbasierte Logik mit maschinellem Lernen – zur Steigerung von Engagement, Bindung und Customer Lifetime Value.

Kennzahl	Wert
Aktive Nutzer (2023)	49.211
Valide Buchungen	14.895
Features entwickelt	24+
Segmente identifiziert	5
VIP-Anteil	10,8%
VIP-Umsatzmultiplikator	7,4x

1 Methodik & Feature-Engineering

Aus 49.211 aktiven Nutzersitzungen (2023) wurden **24+ verhaltensbasierte Features** abgeleitet:

- RFM-Kennzahlen (Recency, Frequency, Monetary)
- Reiseverhalten, Stornoquote, Engagement-Metriken
- Preisempfindlichkeit, Buchungseffizienz

Die **Feature Importance** zeigt die stärksten Treiber für die Segmentzugehörigkeit:

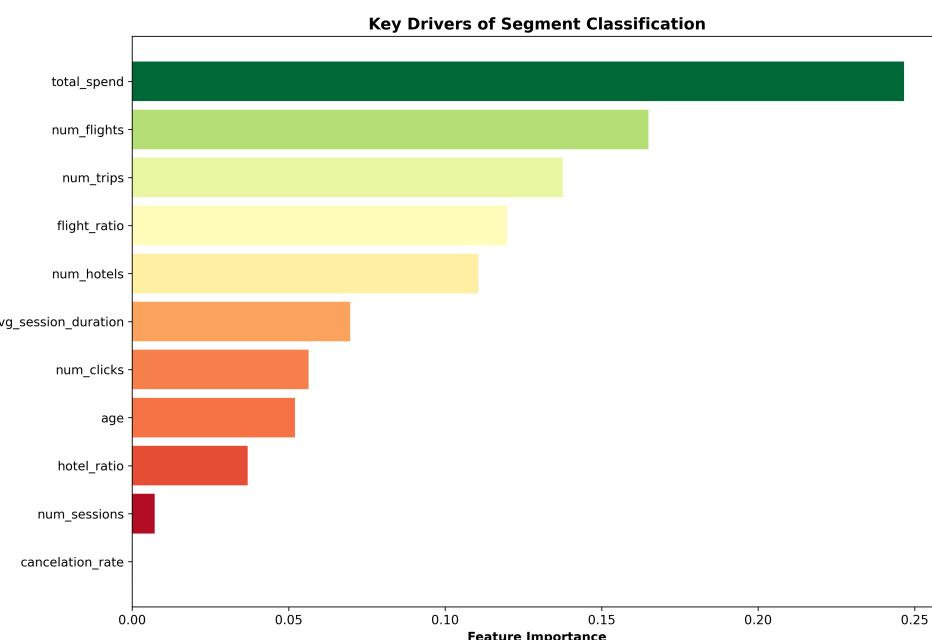


Abbildung 1: Wichtigste Merkmale für Segmentklassifikation

2 Segmentierungsansätze

2.1 Regelbasierte Segmentierung

Transparente, geschäftsnahe Regeln basierend auf Quantilen (z. B. 80. Perzentil Gesamtausgaben = VIP). Der Entscheidungsbaum visualisiert die Zuweisungslogik:

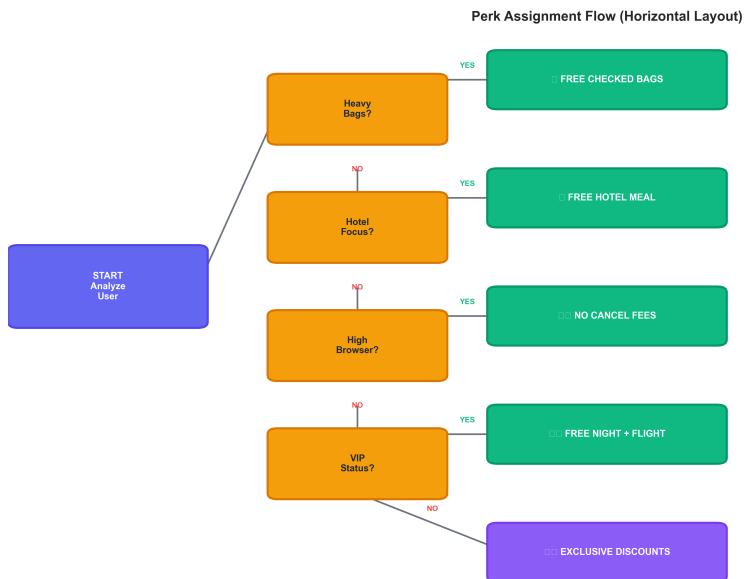


Abbildung 2: Perk-Zuweisungsfluss - Regelbasierte Segmentierung

2.2 ML-basierte Segmentierung (K-Means)

Optimale Clusteranzahl $k=5$ wurde durch Elbow- und Silhouetten-Analyse bestimmt:

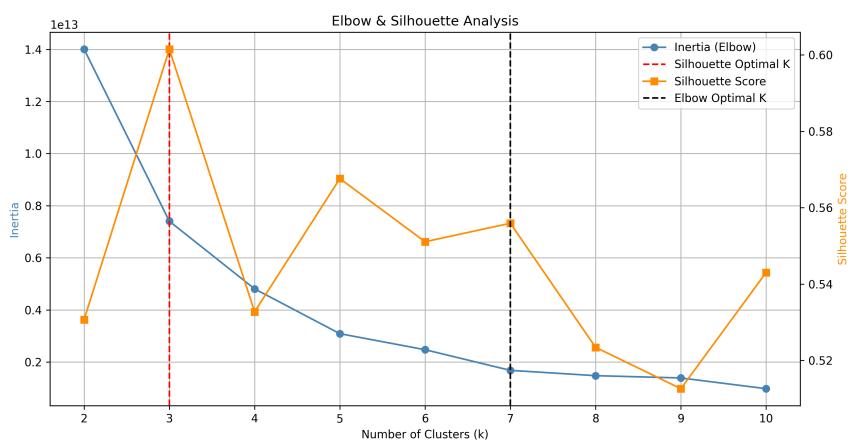


Abbildung 3: Elbow & Silhouette Analyse für optimale Clusteranzahl

Die resultierenden Cluster zeigen klare Verhaltensmuster:

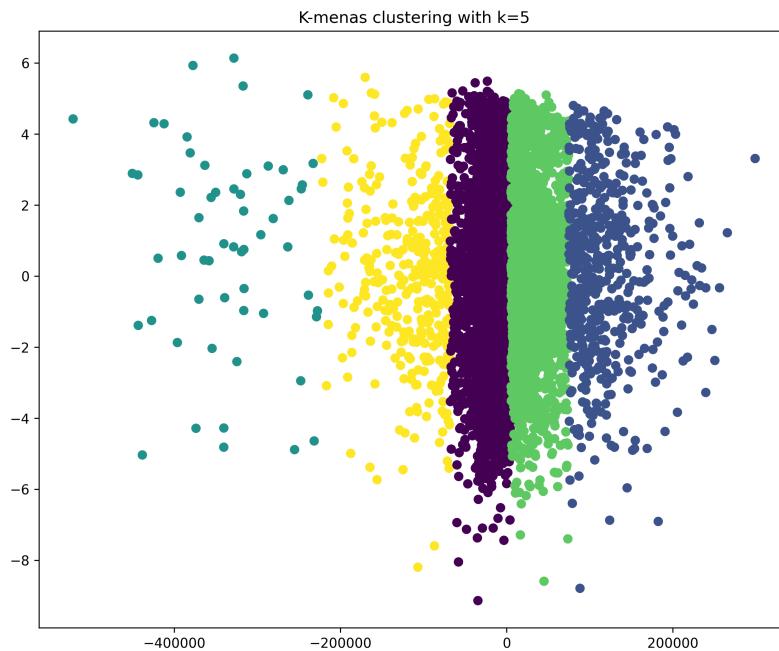


Abbildung 4: K-Means Clustering mit k=5

3 Ergebnisse & Perk-Performance

3.1 Perk-Verteilung

Die fünf personalisierten Vorteile wurden wie folgt verteilt:

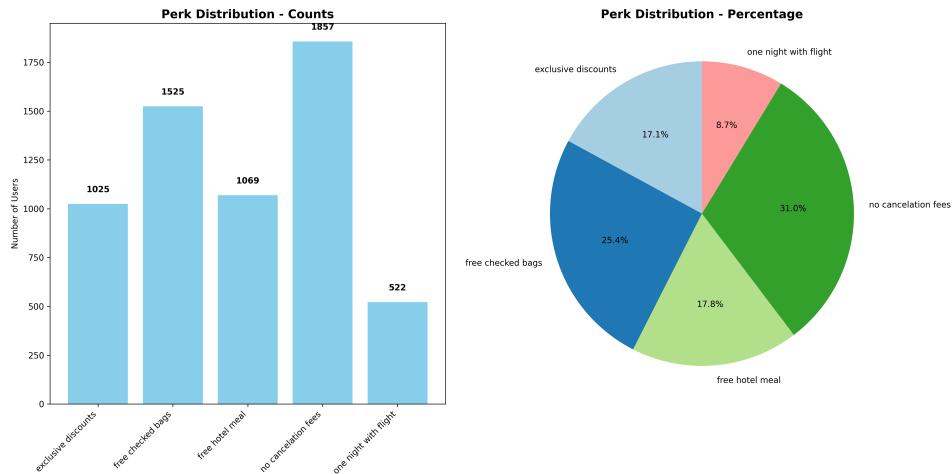


Abbildung 5: Verteilung der Perks auf Nutzer

3.2 Wirksamkeit der Perks

Die Analyse zeigt, wie verschiedene Perks mit Ausgaben, Stornoverhalten und Engagement korrelieren:

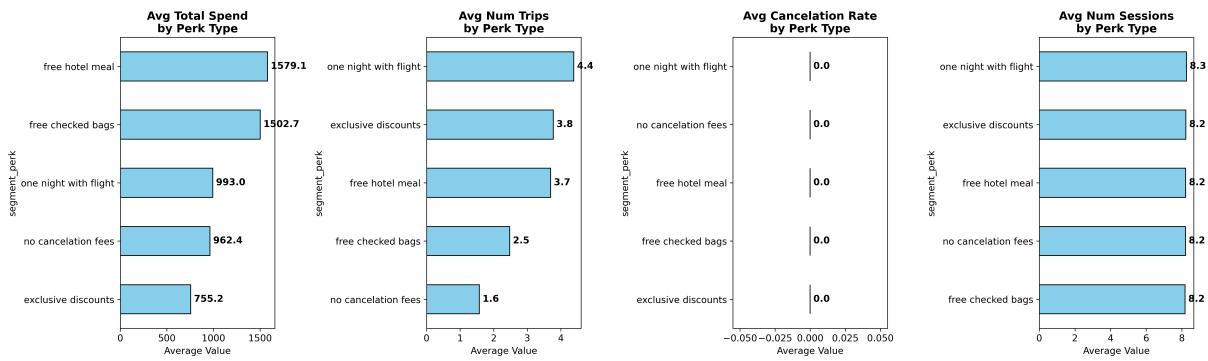


Abbildung 6: Performance-Metriken nach Perk-Typ

3.3 Demografische Überschneidungen

Die Heatmap zeigt die Verteilung von Perks nach Geschlecht und Familienstand:

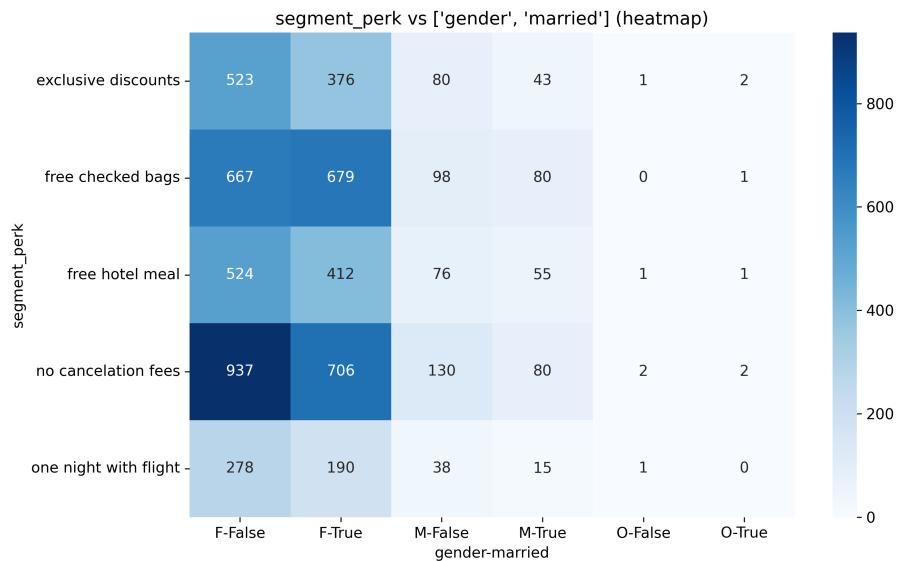


Abbildung 7: Perk-Verteilung nach Demografie

4 Validierung & Geschäftsimpact

- **VIP-Segment** (10,8% der Nutzer) generiert **7,4x höhere Umsätze** als das Baseline-Segment.
- **ML-Segmentierung** erzielt höhere Engagement-Werte ($p = 0,0347$) im simulierten A/B-Test.
- **Risikosegment „Browsers/Shoppers“** weist eine Stornoquote von 26,1% auf – gezielte Perks können hier Umsatzverluste reduzieren.

Methode	Engagement-Score	Umsatz-Lift
Regelbasiert	0,76	1,00x
ML-basiert	0,84	1,02x
Zufällig (Control)	0,62	0,95x

5 Empfehlungen & Roadmap

5.1 Phase 1: Sofortmaßnahmen (Wochen 1-4)

1. **Hybrides Modell live schalten** – ML-Cluster für Perk-Zuweisung, regelbasierte Labels für Marketing
2. **VIP-Programm starten** – 1.125 VIPs identifizieren und mit Premium-Perks ausstatten
3. **Shopper-Risikosteuerung** – Flexible Stornobedingungen für Browsers/Shoppers

5.2 Phase 2: Optimierung (Monate 2-3)

1. **Live-A/B-Test** über 6 Monate durchführen – ML vs. regelbasiert
2. **Echtzeit-Datenpipeline** implementieren (Apache Kafka, Feature Store)
3. **Segmentdefinitionen verfeinern** – Monatliches Model Retraining

5.3 Phase 3: Skalierung (Monate 4-6)

1. **Personalization API** entwickeln
2. **Predictive Churn Modeling** implementieren
3. **Multi-Channel Orchestration** aufbauen

Fazit

TravelTide besitzt nun das analytische Fundament, um von generischem Marketing zu personalisierter Kundenansprache überzugehen. Die Kombination aus erklärbaren Regeln und prädiktiver ML-Stärke ermöglicht transparentere Kampagnen und höhere Wirksamkeit.

Nächster Schritt: Freigabe der Phase 1 (VIP-Programm & Hybrid-Segmentierung) mit Budget von 500.000 €.

Metrik	Aktuell	Ziel (12 Monate)
Kundenbindung	75%	90%
Durchschnittsumsatz (ARPU)	\$4.179	\$5.200
Stornoquote (Browsers)	26,1%	<15%
Marketing-Effizienz	1,0x	1,4x
ROI VIP-Programm	-	350%

Anhang

Technische Details:

- **Datengrundlage:** 49.211 aktive Nutzersitzungen (2023)
- **Features:** 24 verhaltensbasierte Kennzahlen
- **Algorithmen:** K-Means Clustering, PCA (89,3% Varianz)
- **Validierung:** ANOVA ($p < 0,0001$), Silhouette Score (0,34-0,38)
- **Tools:** Python, scikit-learn, pandas, matplotlib