GUEST SEGMENTATION AND PERSONA MODELING USING PYTHON FOR A U.S. HOSPITALITY CHAIN

1. Background

A regional hospitality brand operating 18 boutique hotels across the U.S. needed to revamp its room packages and upselling strategy. Despite a large and diverse guest database, their current personalization efforts were based only on demographics and stay frequency.

The leadership team engaged us to segment their guests using behavioral data such as booking channels, amenity usage, travel companions, and trip purpose. The objective was to create clear guest personas using Python clustering and translate those insights into package offerings, targeted promotions, and loyalty perks.

2. Objective

- To apply unsupervised clustering in Python for guest segmentation based on multi-dimensional booking and usage data
- To extract interpretable personas representing distinct guest types
- To deliver recommendations for customized service bundles and upsell triggers
- To provide scripts and templates for quarterly guest segmentation refresh

3. Data Used

Source: Internal CRM and booking engine logs (2022–2023)

Fields Included:

- Guest_ID, Room_Type, Booking_Channel, Number_of_Stays, Average_Length_of_Stay, In-Room_Dining_Usage, Spa_Usage, WiFi_Speed_Upgrade, Purpose_of_Visit, Travel_Companion_Type
- Size: 42,000 unique guests

4. Methodology

4.1 Data Preparation

- One-hot encoded categorical variables (Purpose_of_Visit, Companion_Type)

- Aggregated amenity usage rates (normalized by number of stays)
- Scaled numeric data using StandardScaler from sklearn

4.2 Clustering Approach

- Applied **K-Means Clustering** (optimal k = 5 determined via Elbow + Silhouette)
- Visualized clusters using PCA to validate separability
- Reviewed and labeled clusters based on behavioral dominance

4.3 Tools Used

- Python Libraries: pandas, scikit-learn, matplotlib, seaborn, numpy

5. Segment Results

Persona Name	% of Guests	Key Traits
Business Minimalists	21%	Solo travelers, short stays, WiFi upgrades, minimal leisure engagement
Luxury Retreaters	17%	Long weekend stays, spa + in-room dining users, premium suite preference
Digital Nomads	15%	Extended stays, booked direct, high WiFi + workspace needs
Social Vacationers	31%	Group bookings, pool/spa users, high amenity consumption
Family Planners	16%	Multi-room bookings, repeat seasonal travelers, loyalty members

6. Interpretation and Strategy

- **Luxury Retreaters** offered curated wellness packages and suite upgrade deals
- **Digital Nomads** received extended stay discounts and co-working offers
- **Social Vacationers** shown event-based bundles (spa + dinner combos)
- **Business Minimalists** targeted with late checkout + laundry packages
- **Family Planners** received loyalty offers tied to school breaks

Each segment also received distinct re-engagement flows and in-app upsell offers via the booking app.

7. Reporting Output

- **Python Jupyter Notebook**:
    - Complete clustering workflow with preprocessing
    - Persona tagging and PCA plots
    - Exported cluster labels with Guest_ID
- **PDF Report (18 pages)**:
    - Persona descriptions and behavior breakdowns
    - Amenity usage patterns and visualization
    - Strategic marketing suggestions per segment
- **Excel Sheet**:
    - Guest_ID with assigned persona
    - Summary tables by geography and loyalty tier
    - Campaign planning matrix based on persona type

8. Business Impact

- Within 3 months of deploying the persona strategy:
    - **Upsell conversion rate** increased by **34%**
    - Package opt-ins increased by **22%**
    - Loyalty enrollment jumped by **19%** among Family Planners and Digital Nomads
    - Segments integrated into CRM and **automated campaigns launched using Mailchimp**