

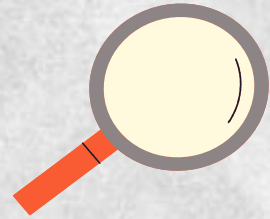


HOTEL BOOKING ANALYSIS

BY VINCENT CHIEN



Source : Kaggle – International Hotel Booking
<https://www.kaggle.com/datasets/alperenmyung/international-hotel-booking-analytics?select=users.csv>



VARIABLE DESCRIPTIONS

hotels

- hotel_id – Hotel identifier
- hotel_name – Name of the hotel
- city – City of the hotel
- country – Country of the hotel
- star_rating – Official star classification
- lat – Latitude of hotel location
- lon – Longitude of hotel location
- cleanliness_base – Baseline cleanliness score
- comfort_base – Baseline comfort score
- facilities_base – Baseline facilities score
- location_base – Baseline location score
- staff_base – Baseline staff service score
- value_for_money_base – Baseline value-for-money score

Reviews

- review_id – Review identifier
- user_id – User identifier
- hotel_id – Hotel identifier
- review_date – Date of review
- score_overall – Overall rating
- score_cleanliness – Cleanliness rating
- score_comfort – Comfort rating
- score_facilities – Facilities rating
- score_location – Location rating
- score_staff – Staff rating
- score_value_for_money – Value-for-money rating
- review_text – Customer comments

Users

- user_id – User identifier
- user_gender – Gender of the user
- country – User's country of residence
- age_group – Age bracket of the user
- traveller_type – Typical travel category (e.g., Solo, Family)
- join_date – Date the user joined the platform



BUSINESS GOAL

- Evaluate Hotel Performance

Identify the top- and bottom-rated hotels based on overall review scores.

- Analyze Satisfaction Drivers

Determine which dimensions (e.g., cleanliness, comfort, facilities, location, staff, value for money) most influence overall satisfaction.

- Explore Visitor Demographics

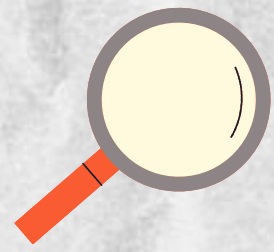
Examine visitor distribution by country and age group to understand key customer segments.

- Understand Key Preferences Across Traveler Types

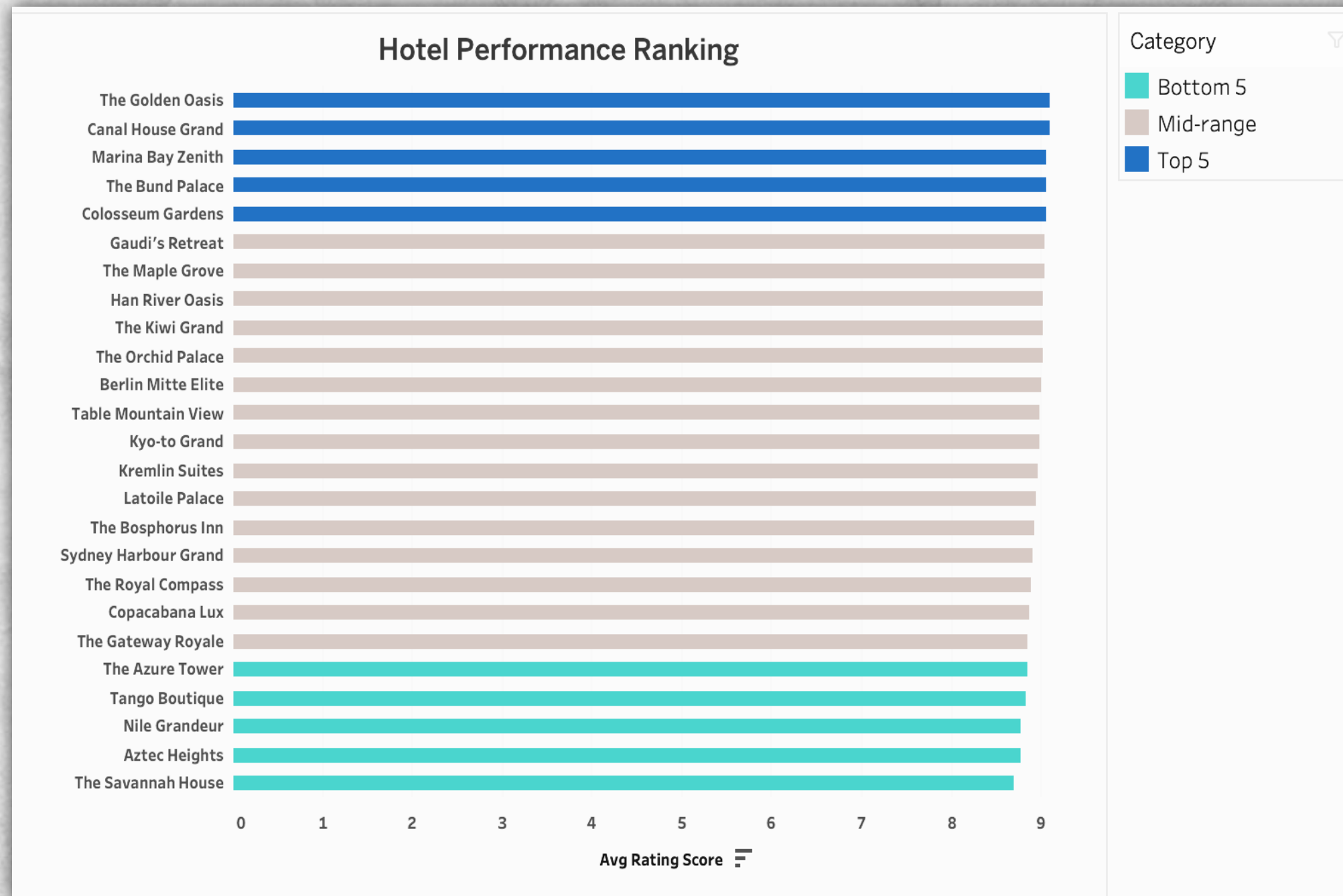
Analyze which satisfaction dimensions are most valued by solo, family, couple, and business travelers.

- Assess Expected vs. Actual Performance

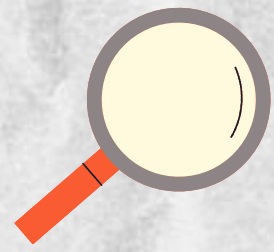
Compare baseline quality scores with actual customer ratings to identify over- and underperforming hotels.



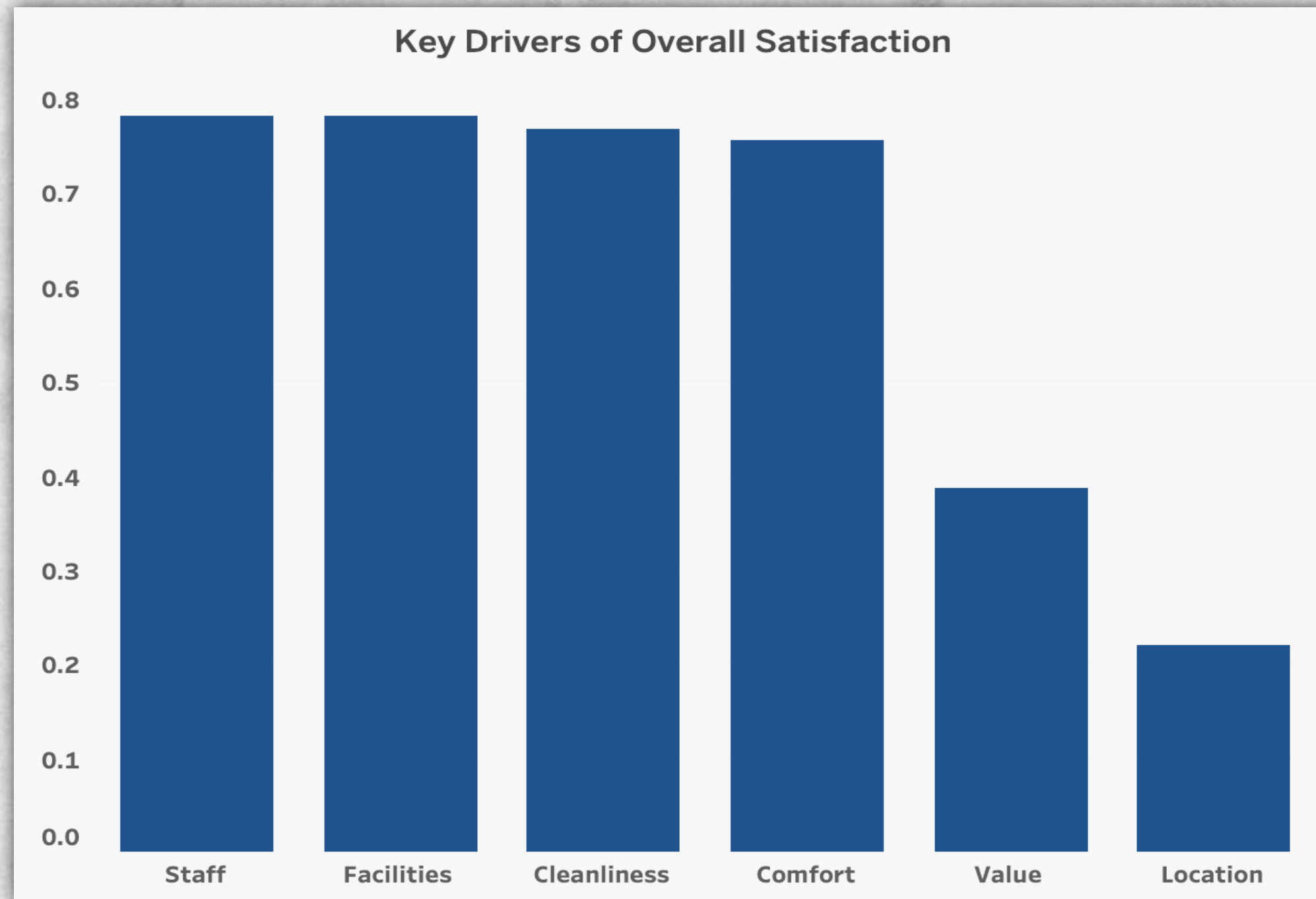
Which hotels perform best and worst based on review scores ?



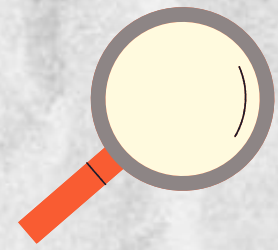
- Top hotels → highest satisfaction (e.g., The Golden Oasis, Canal House Grand)
- Bottom hotels → lowest ratings (e.g., The Savannah House, Aztec Heights)
- Insight → top hotels act as benchmarks; bottom hotels need performance improvement



Which factors drive overall hotel satisfaction the most?



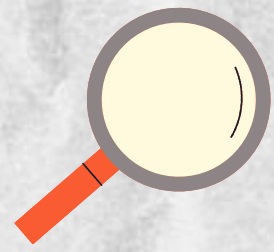
- Staff, facilities, cleanliness, and comfort → strongest drivers of overall satisfaction
- Value and location → weaker influence
- Insight → guests prioritize experience quality (service & comfort) over price or convenience



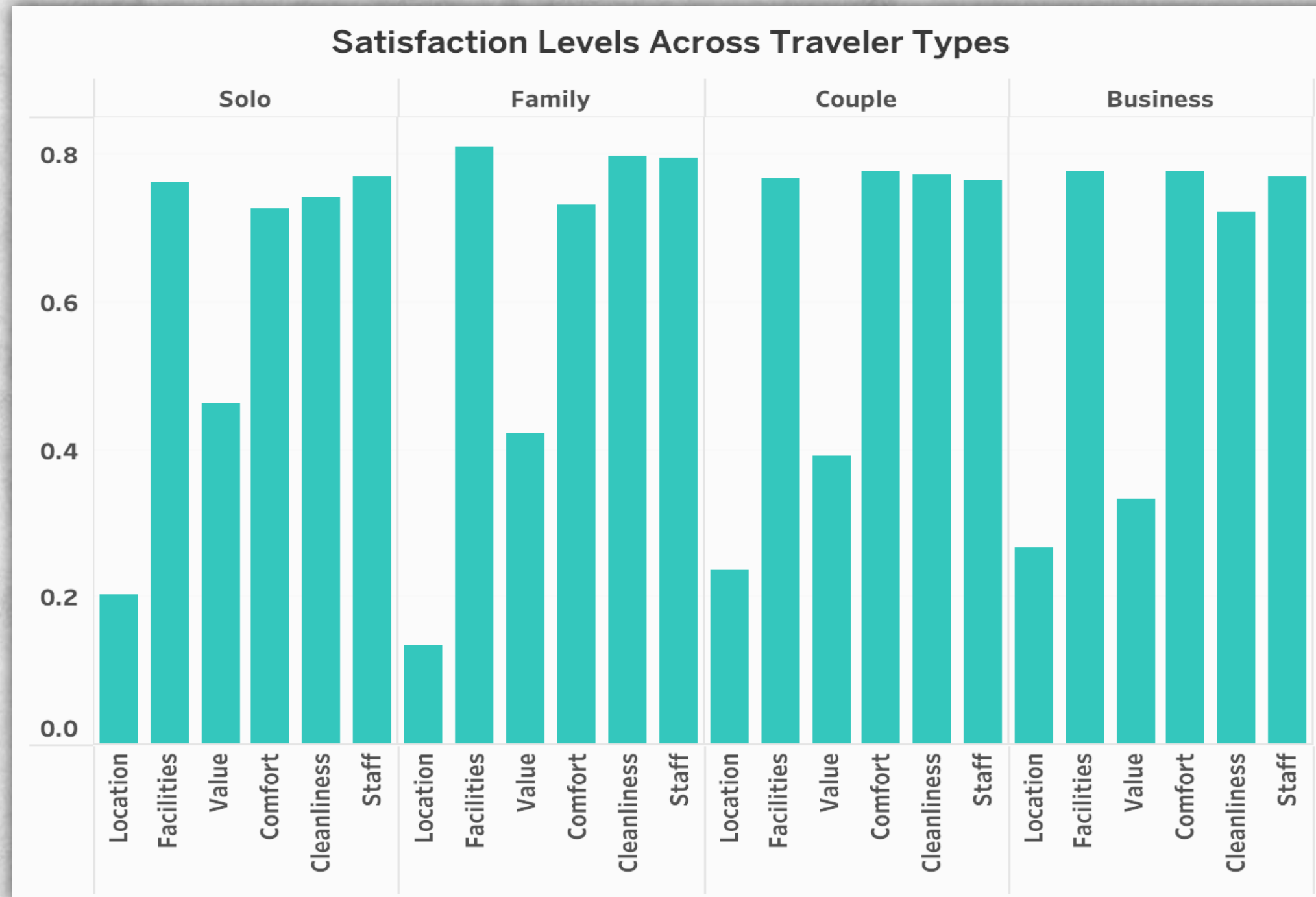
Which regions and age groups dominate hotel visitor demographics?

<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : The Bosphorus Inn City: Istanbul Main visitor country: United States Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : The Orchid Palace City: Bangkok Main visitor country: United Kingdom Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Tango Boutique City: Buenos Aires Main visitor country: United States Dominant age group : 25-34</div></div>
<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Kremlin Suites City: Moscow Main visitor country: United Kingdom Dominant age group : 35-44</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Colosseum Gardens City: Rome Main visitor country: United Kingdom Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Table Mountain View City: Cape Town Main visitor country: United Kingdom Dominant age group : 35-44</div></div>
<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Sydney Harbour Grand City: Sydney Main visitor country: United Kingdom Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Kyo-to Grand City: Tokyo Main visitor country: United Kingdom Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : The Savannah House City: Lagos Main visitor country: United Kingdom Dominant age group : 25-34</div></div>
<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : The Azure Tower City: New York Main visitor country: United States Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : Aztec Heights City: Mexico City Main visitor country: United Kingdom Dominant age group : 25-34</div></div>	<div><div>✓ Keep Only <input type="radio"/> Exclude </div><div>Hotel : The Gateway Royale City: Mumbai Main visitor country: United Kingdom Dominant age group : 35-44</div></div>

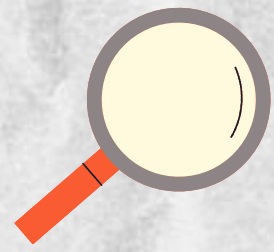
- Dominant regions: UK & USA lead the visitor base
- Age distribution: 25–34 is the primary segment; 35–44 follows
- Insight → tailor service tone & promotions toward young-mid international travelers



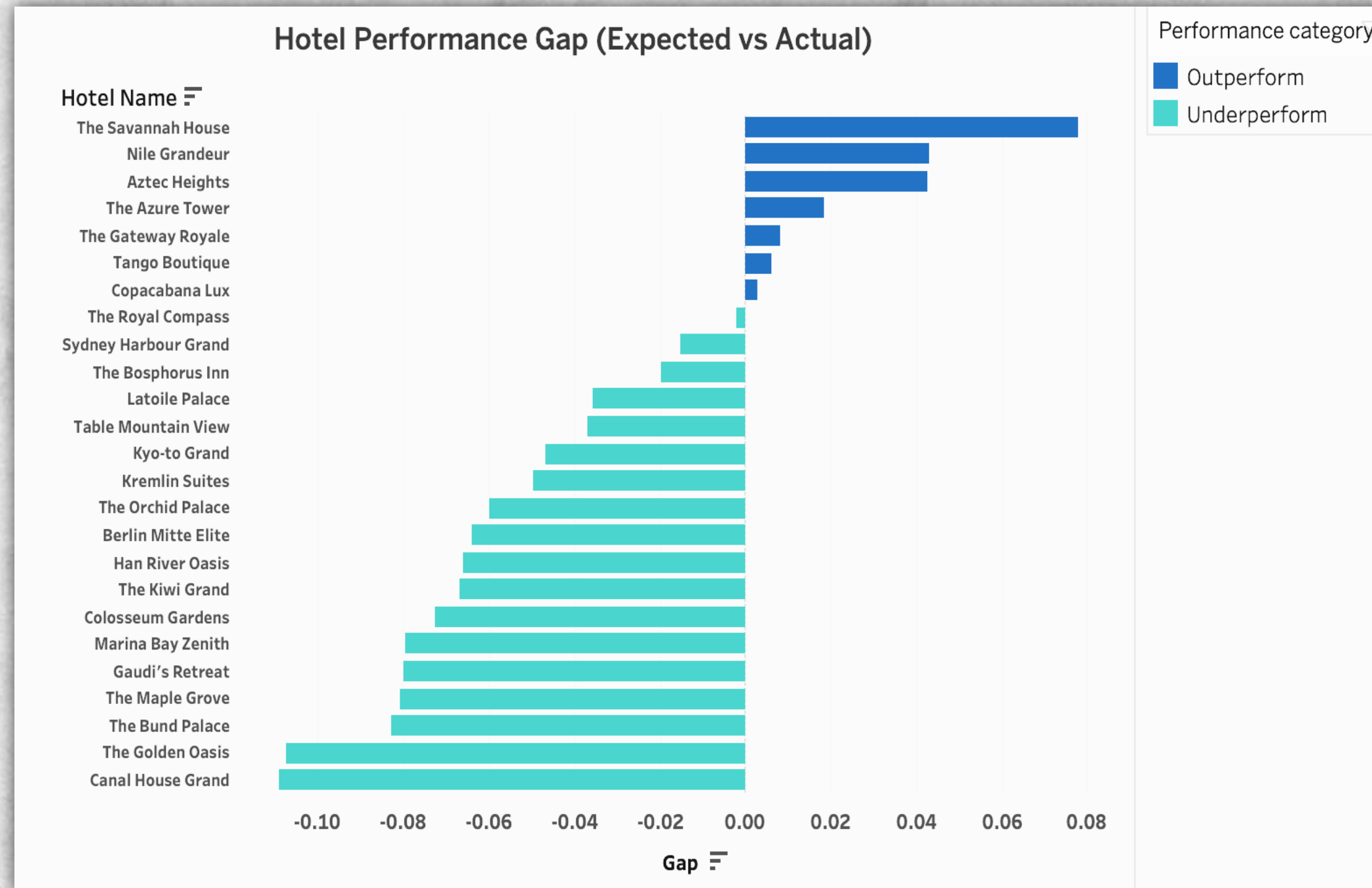
What do different traveler types care about the most?



- Universal → Staff, facilities, comfort & cleanliness score high
- Weak spots → Location & value stay lowest
- By segment →
Solo → staff | Family & Business → facilities | Couples → comfort
- Insight → Enhance experience quality while boosting perceived value & convenience by segment



Which hotels exceed expectations, and which ones fall behind?



- Outperformers (e.g., The Savannah House, Nile Grandeur) exceed expectations
- Underperformers (e.g., Canal House Grand, The Golden Oasis) fall short
- Insight: High-expectation hotels may need refinement; outperformers serve as benchmarks