

Input data

Features: hotel, is_canceled, lead_time, arrival_date_year, arrival_date_month, arrival_date_week_number, arrival_date_day_of_month, stays_in_weekend_nights, stays_in_week_nights, adults, children, babies, meal, country, market_segment, distribution_channel, is_repeated_guest, previous_cancellations, previous_bookings_not_canceled, reserved_room_type, assigned_room_type, booking_changes, deposit_type, company, days_in_waiting_list, customer_type, adr, required_car_parking_spaces, total_of_special_requests, reservation_status, reservation_status_date (total: 31 features)
Meta attributes: agent

Output data

Features: hotel, customer_type, market_segment, distribution_channel, reserved_room_type, lead_time, reservation_status, stays_in_weekend_nights, stays_in_week_nights, total_of_special_requests, adults (total: 11 features)
Target: adr

Removed: 20 (deposit_type, arrival_date_month, agent, required_car_parking_spaces, arrival_date_year, is_repeated_guest, reservation_status_date, previous_bookings_not_canceled, previous_cancellations, assigned_room_type, booking_changes, is_canceled, country, arrival_date_day_of_month, babies, children, meal, company, days_in_waiting_list, arrival_date_week_number)



Number of clusters: 3
Optimization: initialize with KMeans++, 10 re-runs limited to 300 steps

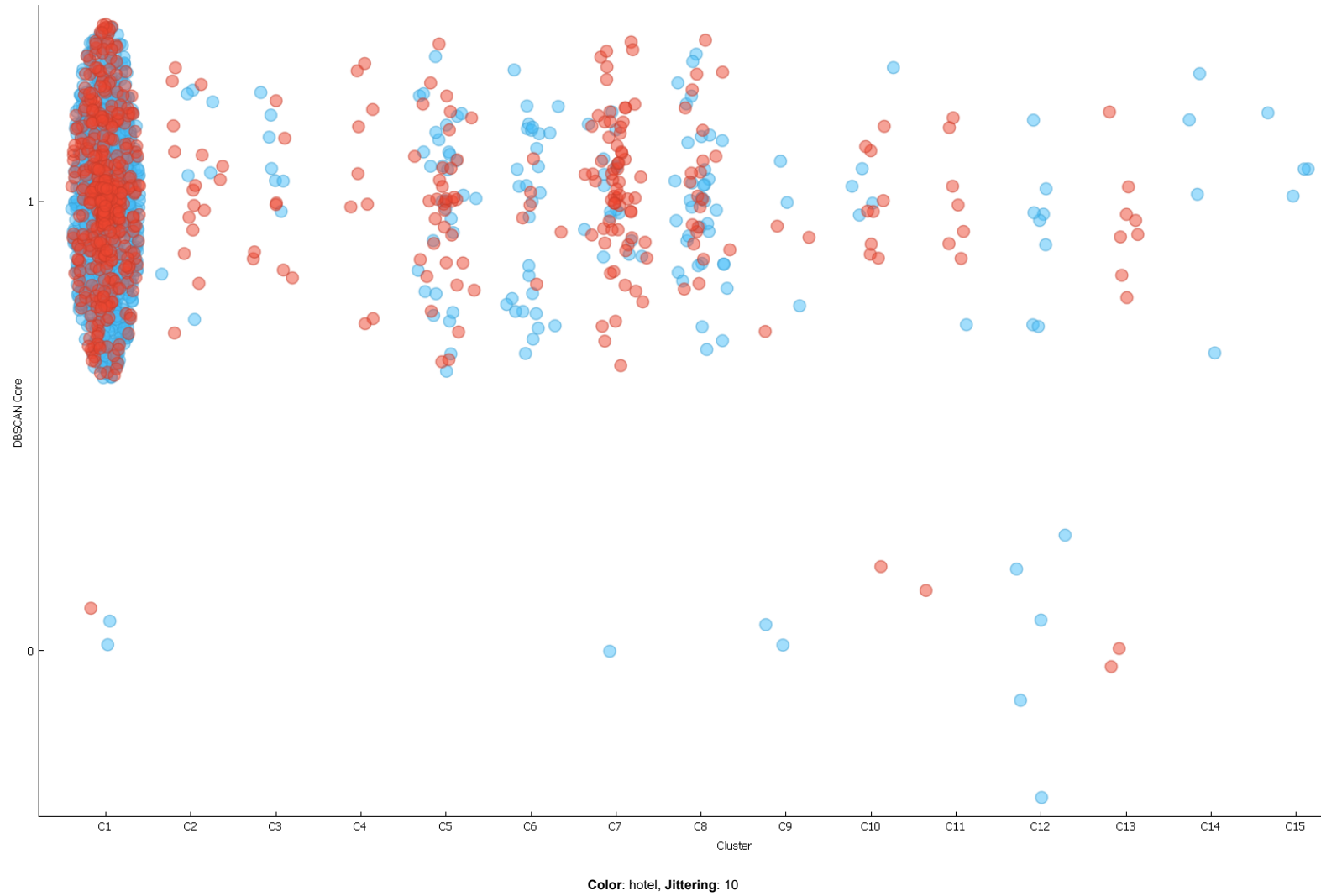
Data

Data instances: 2000
Features: hotel, customer_type, market_segment, distribution_channel, reserved_room_type, lead_time, reservation_status, stays_in_weekend_nights, stays_in_week_nights, total_of_special_requests, adults (total: 11 features)
Target: adr



Normalize data: Yes
PCA preprocessing: Yes, 11 components
Metric: Euclidean
k neighbors: 30
Resolution: 1.0





Test and Score

Sat Jun 04 22, 08:59:01

Settings

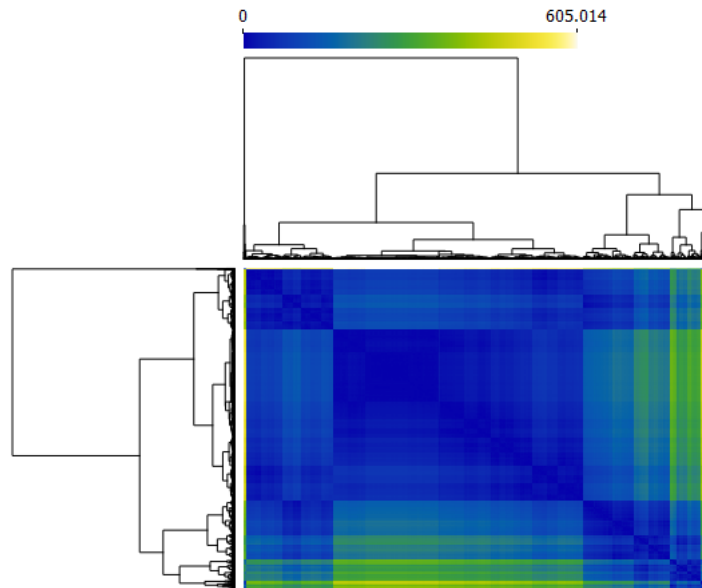
Sampling type: Stratified 3-fold Cross validation
Target class: None, show average over classes

Scores

Model	AUC	CA	F1	Precision	Recall
Logistic Regression	0.9995098072881612	0.987	0.987004222805208	0.9870137211108989	0.987



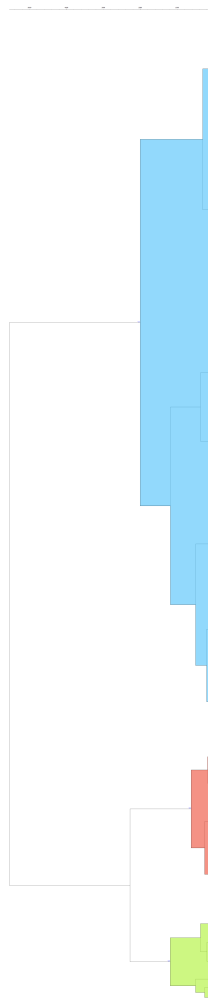
Sorting: clustering
Annotations: none



Hierarchical Clustering

Sat Jun 04 22, 08:59:36

Linkage: ward
Annotation: hotel
Selection: top 3 clusters





Pivot Table

Sat Jun 04 22, 08:59:55

Row feature: Cluster
Column feature: hotel
Value feature: Cluster

		hotel		
Cluster	Count	City Hotel	Resort Hotel	Total
C1	18.0	176.0	194.0	
C2	177.0	4.0	181.0	
C3	148.0	10.0	158.0	
C4	126.0	24.0	150.0	
C5	0.0	144.0	144.0	
C6	92.0	40.0	132.0	
C7	73.0	54.0	127.0	
C8	106.0	4.0	110.0	
C9	3.0	107.0	110.0	
C10	97.0	1.0	98.0	
C11	90.0	0.0	90.0	
C12	66.0	15.0	81.0	
C13	62.0	16.0	78.0	
C14	76.0	0.0	76.0	
C15	0.0	68.0	68.0	
C16	63.0	3.0	66.0	
C17	49.0	13.0	62.0	
C18	45.0	0.0	45.0	
C19	30.0	0.0	30.0	
Total	1321.0	679.0	2000.0	