
ΕΡΓΑΣΙΑ 5

ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

ΑΝΑΦΟΡΑ ΠΑΡΑΔΟΣΗΣ

✚ Παπαθανασίου Αθανασία Μαρία (p3180147)

✚ Ορούτζογλου Αικατερίνη (p3180135)

Για να απαντήσουμε στο ερώτημα εάν η βάση δεδομένων μας είναι σε BCNF πρέπει να γνωρίζουμε τους ορισμούς των κανονικών μορφών 1NF, 2NF και 3NF. Αρχικά, αρκεί ένας πίνακας της βάσης μας να μην είναι σε BCNF προκειμένου να μην είναι όλη η βάση σε BCNF.

➤ 1NF

Ένας πίνακας είναι στην πρώτη κανονική μορφή εάν οι τιμές κάθε στήλης του είναι ατομικές.

➤ 2NF

Ένας πίνακας είναι στη δεύτερη κανονική μορφή εάν ισχύουν οι δύο ακόλουθες προϋποθέσεις: να βρίσκεται στην πρώτη κανονική μορφή και να μην υπάρχει partial dependency, δηλαδή αν $A \rightarrow B$ το A είναι prime attribute και το B είναι non-prime attribute.

➤ 3NF

Ένας πίνακας είναι στη τρίτη κανονική μορφή εάν ισχύουν οι δύο ακόλουθες προϋποθέσεις: να βρίσκεται στη δεύτερη κανονική μορφή και να μην υπάρχει transitive dependency, δηλαδή αν $A \rightarrow B$ το A είναι non-prime attribute και το B είναι non-prime attribute.

➤ BCNF

Ένας πίνακας είναι στην BCNF κανονική μορφή εάν ισχύουν οι δύο ακόλουθες προϋποθέσεις: να βρίσκεται στην τρίτη κανονική μορφή και αν $A \rightarrow B$ μία μη τετριμμένη συναρτησιακή εξάρτηση τότε το A είναι υπερκλειδί. Δηλαδή αν $A \rightarrow B$ το A είναι non-prime attribute και το B είναι prime attribute.

Με βάση τα παραπάνω ελέγχουμε για κάθε πίνακα εάν βρίσκεται σε BCNF μορφή:

Calendar

Ο πίνακας Calendar δεν πληροί την δεύτερη προϋπόθεση της δεύτερης κανονικής μορφής, αφού υπάρχει partial dependency στην συναρτησιακή εξάρτηση $listing_id \rightarrow maximum_nights$. Άρα ο πίνακας δεν είναι σε BCNF.

Amenity

Ο πίνακας Amenity βρίσκεται σε BCNF αφού είναι σε 3NF, δηλαδή δεν υπάρχουν transitive dependencies και το $amenity_id$ είναι superkey στην συναρτησιακή εξάρτηση $amenity_id \rightarrow amenity_name$.

Connection

Ο πίνακας Connection είναι σε BCNF αφού είναι σε 3NF, δηλαδή δεν υπάρχουν transitive dependencies και το $(listing_id, amenity_id)$ είναι υπερκλειδί και ο πίνακας δεν περιέχει άλλες στήλες, οπότε δεν υπάρχουν κακές συναρτησιακές εξαρτήσεις.

Geolocation

Ο πίνακας Geolocation δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στην συναρτησιακή εξάρτηση $type \rightarrow geometry_type$. Οπότε ο πίνακας δεν είναι σε BCNF.

Host

Ο πίνακας Host δε βρίσκεται στην πρώτη κανονική μορφή, αφού οι τιμές της στήλης verifications δεν είναι ατομικές. Άρα ο πίνακας δεν είναι σε BCNF.

Listing_Summary

Ο πίνακας Listing_Summary δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στην συναρτησιακή εξάρτηση $host_id \rightarrow \{host_name, neighbourhood, minimum_nights\}$. Άρα ο πίνακας δεν είναι σε BCNF.

Location

Ο πίνακας Location δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στις συναρτησιακές εξαρτήσεις $neighbourhood_cleansed \rightarrow neighbourhood$ και $country \rightarrow \{country_code, market\}$. Άρα ο πίνακας δεν είναι σε BCNF.

Neighbourhood

Ο πίνακας Neighbourhood βρίσκεται σε BCNF αφού είναι σε 3NF, δηλαδή δεν υπάρχουν transitive dependencies και το neighbourhood είναι superkey στην συναρτησιακή εξάρτηση $neighbourhood \rightarrow neighbourhood_group$.

Price

Ο πίνακας Price δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στις συναρτησιακές εξαρτήσεις {minimum_nights, maximum_nights}->{ minimum_minimum_nights, maximum_minimum_nights, minimum_maximum_nights, maximum_maximum_nights, minimum_nights_avg_ntm, maximum_nights_avg_ntm}. Άρα ο πίνακας δεν είναι σε BCNF.

Review

Ο πίνακας Review δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στις συναρτησιακές εξαρτήσεις {listing_id, reviewer_id}->{reviewer_name, date, comments}. Άρα ο πίνακας δεν είναι σε BCNF.

Review_Summary

Ο πίνακας Review_Summary είναι σε BCNF αφού είναι σε 3NF, δηλαδή δεν υπάρχουν transitive dependencies και το (listing_id,date) είναι υπερκλειδί και ο πίνακας δεν περιέχει άλλες στήλες, οπότε δεν υπάρχουν κακές συναρτησιακές εξαρτήσεις.

Room

Ο πίνακας Room είναι σε BCNF αφού είναι σε 3NF, δηλαδή δεν υπάρχουν transitive dependencies και το listing_id είναι υπερκλειδί που καθορίζει όλα τα υπόλοιπα γνωρίσματα.

Listing

Ο πίνακας Listing δεν πληροί την δεύτερη προϋπόθεση της τρίτης κανονικής μορφής αφού υπάρχει transitive dependency στις συναρτησιακές εξαρτήσεις {number_of_reviews , first_review, last_review}->{ number_of_reviews_ltm, review_scores_rating, review_scores_accuracy, review_scores_cleanliness, review_scores_checkin, review_scores_communication, review_scores_location, review_scores_value} και host_id->{property_type,room_type ,calculated_host_listings_count_entire_homes,calculated_host_listings_count_private_rooms , calculated_host_listings_count_shared_rooms}. Άρα ο πίνακας δεν είναι σε BCNF.

