

Εργασία 3η: Neo4j

A. Σχεδιάστε ένα γράφο ιδιοτήτων για την παρακάτω σχεσιακή βάση και αναφέρετε ποιοι περιορισμοί που επιβάλλονται στο σχεσιακό μοντέλο δεν ισχύουν στο γράφο ιδιοτήτων που σχεδιάσατε (αν υπάρχουν τέτοιοι περιορισμοί). Υπογραμμισμένα είναι τα κύρια κλειδιά των πινάκων και με πλάγια γραφή και ονοματολογία πεδίο-πίνακας τα ξένα κλειδιά.

Φοιτητής(AEM, όνομα)

Μάθημα(κωδικός, τίτλος, διδακτικές-μονάδες)

Διδάσκουσα(AΦM, όνομα, διεύθυνση)

Αμφιθέατρο(αριθμός, θέσεις)

Επιτηρήτρια(κωδικός, όνομα, τηλέφωνο)

Τμήμα(κωδ-μάθημα, α/α, AΦM, αριθ-αμφιθέατρο, ημέρα, ώρα)

κωδ-μάθημα ξένο κλειδί αναφέρεται στο Μάθημα(κωδ)

AΦM ξένο κλειδί αναφέρεται στο Διδάσκουσα(AΦM)

αριθ-αμφιθέατρο ξένο κλειδί αναφέρεται στο Αμφιθέατρο(αριθμός)

Ανήκει(AEM, κωδ-μάθημα-τμήμα, α/α-τμήμα)

AEM ξένο κλειδί αναφέρεται στο Φοιτητής(AEM)

(κωδ-μάθημα-τμήμα, α/α-τμήμα) ξένο κλειδί αναφέρεται στο Τμήμα(κωδ-μάθημα, α/α)

Εξέταση(κωδ-μάθημα, εξεταστική, ώρα, ημερομηνία)

κωδ-μάθημα ξένο κλειδί αναφέρεται στο Μάθημα(κωδ)

Φιλοξενείται(κωδ-μάθημα-εξέταση, εξεταστική-εξέταση, αριθμ-αμφιθέατρο, κωδ-επιτηρήτρια)

(κωδ-μάθημα-εξέταση, εξεταστική-εξέταση) ξένο κλειδί αναφέρεται στο Εξέταση(κωδ-μάθημα, εξεταστική)

αριθ-αμφιθέατρο ξένο κλειδί αναφέρεται στο Αμφιθέατρο(αριθμός)

κωδ-επιτηρήτρια ξένο κλειδί αναφέρεται στο Επιτηρήτρια(κωδ)

Βαθμολογείται(AEM, κωδ-μάθημα, εξεταστική, βαθμός)

AEM ξένο κλειδί αναφέρεται στο Φοιτητής(AEM)

(κωδ-μάθημα-εξέταση, εξεταστική-εξέταση) ξένο κλειδί αναφέρεται στο Εξέταση(κωδ-μάθημα, εξεταστική)

B. Σχεδιάστε ένα γράφο ιδιοτήτων για ένα ταξιδιωτικό γραφείο που παρακολουθεί τους πελάτες του και τις προτιμήσεις τους, και αναφέρετε ποιοι από τους περιορισμούς που περιγράφονται δεν ισχύουν στο γράφο ιδιοτήτων που σχεδιάσατε (αν υπάρχουν τέτοιοι περιορισμοί).

- Κάθε πελάτης προσδιορίζεται από το μοναδικό του email και έχει επίσης ένα όνομα και έναν τηλεφωνικό αριθμό.
- Οι αεροπορικές εταιρίες ταυτοποιούνται από τα URL τους, αλλά έχουν επίσης ένα αεροδρόμιο έδρα και ένα μοναδικό όνομα.
- Οι πελάτες μπορεί να έχουν λογαριασμούς τακτικών πελατών (frequent flyer) με διάφορες εταιρίες, οπότε και διατηρούμε την πληροφορία για κάθε τέτοια αεροπορική εταιρία τον αριθμό μέλους τους σε αυτήν ο οποίος είναι μοναδικός για κάθε εταιρία καθώς και τους αντίστοιχους πόντους.
- Οι προορισμοί διακοπών έχουν ένα κωδικό και ένα όνομα. Κάθε προορισμός βρίσκεται σε μοναδική χώρα, για την οποία διατηρούμε το όνομα και το νόμισμά της. Ωστόσο δεν έχει κάθε χώρα στη λίστα μας κάποιον προορισμό διακοπών.
- Για κάθε πελάτη και προορισμό, υπάρχει μια μοναδική εταιρία που ο πελάτης προτιμά για αυτόν τον προορισμό. Πολλοί πελάτες μπορεί να προτιμούν την ίδια αεροπορική εταιρία για έναν δεδομένο προορισμό, και ένας πελάτης μπορεί να προτιμά την ίδια εταιρία για πολλούς προορισμούς.

Γ. Θα εργαστείτε στο Neo4j sandbox με τη βάση δεδομένων: StackOverflow.

(<https://neo4j.com/sandbox/>) Επιλέγεται New Project και στα Pre-built data θα βρείτε το Stack Overflow.

Καλέστε το “call db.schema.visualization()” για να δείτε και να μελετήσετε το σχήμα του γράφου. Αφού μελετήσετε τη δομή του γράψτε τον κώδικα Cypher για τα ακόλουθα ερωτήματα (σε παρένθεση το πλήθος των αποτελεσμάτων).

Σημείωση: το sandbox πιθανόν σβήνει τη βάση σας μετά από τρεις μέρες. Μπορείτε να την ξαναδημιουργήσετε χωρίς πρόβλημα.

1. Βρείτε τις ερωτήσεις που έχουν tag και “cypher” και “indexing”. Εμφανίστε tags και ερωτήσεις. (5 nodes)
2. Βρείτε τους χρήστες που έχουν κάνει ερώτηση που δεν έχει καμία απάντηση αλλά έχουν σχόλια. Εμφανίστε τα display name των χρηστών. (208 records)
3. Βρείτε τον χρήστη του οποίου οι ερωτήσεις έχουν χρησιμοποιήσει συνολικά τα περισσότερα διαφορετικά tags. Εμφανίστε το display name του χρήστη και το πλήθος των tags. (1 record)
4. Βρείτε την ερώτηση με τα περισσότερα views (μεγαλύτερο view_count). Εμφανίστε τον κόμβο. (1 node)
5. Βρείτε τους χρήστες που ούτε έχουν απαντήσει, ούτε έχουν σχολιάσει δική τους ερώτηση. Εμφανίστε τα display name των χρηστών. (811 records)
6. Για τις ερωτήσεις που έχουν tags, οι ερωτήσεις που έχουν τουλάχιστον 10 σχόλια. Να εμφανιστεί ο τίτλος της ερώτησης, το πλήθος των σχολίων και λίστα με τα tags τους. (5 records)
7. Βρείτε ζεύγη από tags που εμφανίζονται μαζί σε πάνω από 10 ερωτήσεις. Εμφανίστε το ζεύγος των ονομάτων των tags και το πλήθος των εμφανίσεων τους διατεταγμένα σε φθίνουσα σειρά με βάση το πλήθος εμφανίσεων. (55 records)
8. Βρείτε τα μονοπάτια μεταξύ όλων των σχολίων που έχουν σκορ μεγαλύτερο από 1 με μήκος το πολύ 5. Εμφανίστε τα μονοπάτια. (34 nodes)
9. Βρείτε το συντομότερο μονοπάτι μεταξύ των tags που περιέχουν τη λέξη neo4j αλλά δεν είναι “neo4j”. Εμφανίστε ζεύγη tag names και το μήκος του μονοπατιού που τα ενώνει. (136 records)
10. Βρείτε τους χρήστες που και έχουν ρωτήσει κάτι, και έχουν απαντήσει σε κάτι και έχουν σχολιάσει κάτι. Εμφανίστε το display_name του χρήστη. (66 records).

Για τη σχεδίαση των γράφων χρησιμοποιείτε το arrow tool (<https://www.apcjones.com/arrows/>) και συγκεκριμένα επιλέξτε το Arrows.app (Κουμπί Try Arrows App).

Για τα Α και Β παραδώστε ένα pdf με τους γράφους ως εικόνα (png) και όποιον σχολιασμό έχετε.

Για το Γ παραδώστε ένα txt αρχείο με τον κώδικα των ερωτημάτων (να φαίνεται ο αριθμός του ερωτήματος στο οποίο αντιστοιχεί η κάθε απάντησή σας).