



ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ (4^ο ΕΞ.)

ΕΡΓΑΣΙΑ ΜΑΘΗΜΑΤΟΣ

(ομάδες των 2-3 ατόμων)

Εισαγωγή

Έστω η ΒΔ μιας αεροπορικής εταιρείας, στην οποία διατηρούνται πληροφορίες σχετικά με τις έννοιες κράτηση (book), εισιτήριο (ticket), κάρτα επιβίβασης (boarding pass), αεροδρόμιο (airport), πτήση (flight) κλπ. Πιο συγκεκριμένα:

- **Κράτηση.** Για να πετάξουν με την αεροπορική μας εταιρεία, οι επιβάτες κλείνουν τα απαιτούμενα εισιτήρια εκ των προτέρων (book_date, το οποίο δεν πρέπει να είναι νωρίτερα από ένα μήνα πριν από την πτήση). Η κράτηση αναγνωρίζεται από τον αριθμό της (book_ref, ένας συνδυασμός έξι θέσεων από γράμματα και ψηφία) και μπορεί να περιλαμβάνει πολλούς επιβάτες, με ξεχωριστό όμως εισιτήριο για τον καθένα. Για κάθε κράτηση θέλουμε να κρατάμε τη συνολική αξία των εισιτηρίων.
- **Εισιτήριο.** Ένα εισιτήριο έχει έναν μοναδικό αριθμό (ticket_no), ο οποίος αποτελείται από 13 ψηφία. Το εισιτήριο περιέχει τον αριθμό του εγγράφου ταυτότητας του επιβάτη (passenger_id), το όνομα και το επώνυμό του (passenger_name) καθώς και τα στοιχεία επικοινωνίας του (contact_data). Για λόγους απλότητας, ας υποθέσουμε ότι όλοι οι επιβάτες είναι μοναδικοί. Ένα εισιτήριο περιέχει ένα ή περισσότερα τμήματα πτήσεων, το καθένα με την τιμή του (amount) και την κατηγορία ταξιδιού (fare – επιλογή από Economy / Business / First class). Είναι σύνηθες να συμπεριλαμβάνονται περισσότερα από ένα τμήματα πτήσεων στο ίδιο εισιτήριο, π.χ. όταν αφορά σε ταξίδι με ενδιάμεσους σταθμούς ή πρόκειται για εισιτήριο μετ' επιστροφής. Παραδοχή: όλα τα εισιτήρια μιας συγκεκριμένης κράτησης περιλαμβάνουν τα ίδια τμήματα πτήσης.

- Πτήση:** Μια πτήση πηγαίνει από ένα αεροδρόμιο σε ένα άλλο. Ο αριθμός πτήσης αφορά το ζεύγος αεροδρόμιο αναχώρησης (departure airport) και αεροδρόμιο άφιξης (arrival airport) – προσοχή: δεν επιτρέπεται τα δύο αεροδρόμια μιας πτήσης να ταυτίζονται – και για να ταυτοποιηθεί η συγκεκριμένη πτήση προστίθεται και η ημερομηνία αναχώρησης (departure date). Συνοδευτική πληροφορία αποτελεί το μοντέλο αεροσκάφους που εκτελεί την πτήση, η χιλιομετρική απόσταση (distance), οι προγραμματισμένες ώρες αναχώρησης (scheduled_departure_time) και άφιξης (scheduled_arrival_time), η προγραμματισμένη διάρκεια της πτήσης (scheduled_duration, προκύπτει ως η αφαίρεση των δύο ωρών, λαμβάνοντας υπόψη τις ζώνες ώρας των δύο αεροδρομίων), οι πραγματικές ώρες αναχώρησης (actual_departure_time) και άφιξης (actual_arrival_time), οι οποίες πολύ συχνά διαφέρουν από τις προγραμματισμένες, και, τέλος, η κατάσταση πτήσης (flight_status) η οποία μπορεί να λάβει μία από τις ακόλουθες τιμές:
 - Scheduled (από 1 μήνα έως 24 ώρες πριν την προγραμματισμένη ώρα αναχώρησης): η πτήση είναι διαθέσιμη για κράτηση αλλά δεν έχει ανοίξει ακόμη το check-in. Πριν από αυτό το χρονικό διάστημα του 1 μηνός δεν μπορεί να γίνει κράτηση για την πτήση.
 - OnTime (από 24 ώρες πριν την προγραμματισμένη ώρα αναχώρησης έως την προγραμματισμένη ώρα αναχώρησης): η πτήση είναι ανοιχτή για check-in.
 - Delayed (μετά την προγραμματισμένη και πριν την πραγματική ώρα αναχώρησης, εφόσον η πτήση δεν έχει ακυρωθεί): η πτήση είναι ανοιχτή για check-in αλλά υπάρχει καθυστέρηση.
 - Departed (μετά την πραγματική ώρα αναχώρησης και πριν την πραγματική ώρα άφιξης): Το αεροσκάφος έχει αναχωρήσει και βρίσκεται εν πτήση.
 - Arrived (μετά την πραγματική ώρα άφιξης): Το αεροσκάφος έχει φτάσει στον προορισμό του.
 - Cancelled: Η πτήση ακυρώθηκε.
- Αεροδρόμιο.** Ένα αεροδρόμιο αναγνωρίζεται από έναν κωδικό τριών γραμμάτων (code) που συνοδεύεται από το όνομα (name), την πόλη (city) και τη ζώνη ώρας (timezone) του αεροδρομίου.
- Αεροσκάφος.** Κάθε μοντέλο αεροσκάφους αναγνωρίζεται από έναν τριψήφιο κωδικό (aircraft_code) και συνοδευτική πληροφορία αποτελεί το όνομα του μοντέλου, η χωρητικότητά του ως πλήθος θέσεων (capacity) και η εμβέλεια (range), δηλαδή η μέγιστη χιλιομετρική απόσταση που μπορεί να καλύψει. Προσοχή: δεν επιτρέπεται μια πτήση να ανατεθεί σε μοντέλο αεροσκάφους με εμβέλεια μικρότερη από την απόσταση της πτήσης.
- Κάρτα επιβίβασης.** Κατά τον έλεγχο πτήσης, ο επιβάτης λαμβάνει μια κάρτα επιβίβασης (διαδικασία check-in), στην οποία αναγράφονται το όνομά του, η πτήση και ο αριθμός της θέσης του (seat_no). Οι κάρτες επιβίβασης λαμβάνουν διαδοχικούς αριθμούς (boarding_no), με τη σειρά των check-in για την πτήση (ο αριθμός αυτός είναι μοναδικός

μόνο στο πλαίσιο μιας συγκεκριμένης πτήσης). Προσοχή: (α) ο επιβάτης μπορεί να λάβει κάρτα επιβίβασης μόνο εάν η πτήση περιλαμβάνεται στο εισιτήριό του και μόνο εάν η πτήση είναι σε κατάσταση OnTime ή Delayed, (β) ο συνδυασμός πτήσης-θέσης πρέπει να είναι μοναδικός για να αποφευχθεί η έκδοση δύο καρτών επιβίβασης για την ίδια θέση, (γ) δεν επιτρέπεται να εκδοθούν περισσότερες κάρτες επιβίβασης από όσες θέσεις διαθέτει το μοντέλο αεροσκάφους που εκτελεί την πτήση.

Ερώτημα 1 (30%). Σχεσιακή Βάση Δεδομένων

- a. Με βάση τα παραπάνω στοιχεία, σχεδιάστε το σχεσιακό σχήμα της ΒΔ, να το υλοποιήσετε (εντολές CREATE TABLE) στο ΣΔΒΔ PostgreSQL και φορτώστε με δεδομένα τους πίνακες. Επιπλέον, καλείστε να τεκμηριώσετε τους περιορισμούς ακεραιότητας των πινάκων (και να δηλώσετε τυχόν περιορισμούς που προκύπτουν από την εκφώνηση αλλά δεν μπορέσατε να υποστηρίξετε μέσα από τους περιορισμούς ακεραιότητας των πινάκων). Το παραδοτέο του υποερωτήματος είναι το σχεσιακό σχήμα της ΒΔ, οι εντολές CREATE TABLE και τα αρχεία τα οποία θα εισάγετε στους πίνακες. Οδηγία: για την ευκολότερη παραγωγή **αληθοφανών** δεδομένων προτείνεται να χρησιμοποιήσετε κάποιο εργαλείο παραγωγής δεδομένων (data generator) (π.χ. www.mockaroo.com, <https://faker.readthedocs.io/en/master/>, <https://devskiller.github.io/jfairly/>).
- b. Εφαρμόστε τη θεωρία της κανονικοποίησης πάνω στο σχεσιακό σχήμα της ΒΔ που σχεδιάσατε και ελέγξτε τον κάθε πίνακα εάν ακολουθεί την BCNF. Σε αντίθετη περίπτωση, αποσυνθέστε τους προβληματικούς πίνακες ώστε όλη η ΒΔ να είναι σε BCNF.
- c. Πάνω στο τελικό σχήμα της ΒΔ υλοποιήστε 2 προβολές/όψεις (views):
 - **Flights_View**. Μια προβολή που θα αφορά μια συγκεκριμένη ημερομηνία (π.χ. 30/5/2022) και θα περιλαμβάνει τις «δυναμικές» πληροφορίες των πτήσεων εκείνης της ημέρας: αριθμός πτήσης, αεροδρόμιο (κωδικός – όνομα) και πόλη αναχώρησης, αεροδρόμιο (κωδικός – όνομα) και πόλη άφιξης, τοπική ημερομηνία και ώρα αναχώρησης (προγραμματισμένη & πραγματική), τοπική ημερομηνία και ώρα άφιξης (προγραμματισμένη & πραγματική), διάρκεια πτήσης (προγραμματισμένη & πραγματική).
 - **Routes_View**. Μια προβολή που θα αφορά μια συγκεκριμένη εβδομάδα (π.χ. Δε. 30/5/2022 – Κυ. 5/6/2022) και θα περιλαμβάνει τις «στατικές» πληροφορίες των πτήσεων εκείνης της εβδομάδας: αριθμός πτήσης, αεροδρόμιο (κωδικός – όνομα) και πόλη αναχώρησης, αεροδρόμιο (κωδικός – όνομα) και πόλη άφιξης, μοντέλο αεροσκάφους, προγραμματισμένη διάρκεια πτήσης και ένα αλφαριθμητικό «days_of_week» το οποίο εμφανίζει τις ημέρες της εβδομάδας κατά τις οποίες εκτελείται η συγκεκριμένη πτήση (π.χ. το αλφαριθμητικό "Mo -- We -- -- Sa --" αντιστοιχεί σε μια πτήση που εκτελείται Δευτέρα, Τετάρτη και Σάββατο).

Ερώτημα 2 (30%). Εκτελέστε τις παρακάτω ερωτήσεις (queries) στη ΒΔ (εντολές SELECT).

- a. Ποιος ταξίδεψε με μια συγκεκριμένη πτήση (π.χ. PG0404) χθες στη θέση 1A και πότε έγινε η κράτηση του εισιτηρίου; Η χθεσινή ημερομηνία θα υπολογίζεται με βάση την ημερομηνία εκτέλεσης του ερωτήματος.
- b. Πόσες θέσεις παρέμειναν ελεύθερες στην ανωτέρω πτήση;
- c. Ποιες πτήσεις είχαν τις μεγαλύτερες καθυστερήσεις μέσα στο 2022; Εμφανίστε τη λίστα των 5 πρώτων.
- d. Βρείτε τους 5 πιο συχνούς ταξιδιώτες (frequent travelers) μέσα στο 2022, δηλαδή αυτούς που έκαναν τα περισσότερα χιλιόμετρα πτήσεων.
- e. Βρείτε τους 5 πιο δημοφιλείς προορισμούς μέσα στο 2022, δηλαδή τις πόλεις προς τις οποίες ταξίδεψαν οι περισσότεροι επιβάτες.
- f. Βρείτε τους πιο «γρήγορους» επιβάτες, δηλαδή αυτούς που έκαναν check-in πρώτοι για όλες τις πτήσεις τους. Λάβετε υπόψη μόνο τους επιβάτες που έκαναν τουλάχιστον δύο πτήσεις.

Ερώτημα 3 (20%). Υλοποίηση triggers, cursors

- a. Φτιάξτε έναν trigger ο οποίος κρατά/γεμίζει ένα πίνακα-ιστορικό. Ο trigger ενεργοποιείται στις μεταβολές διαγραφή και ενημέρωση του κύριου πίνακα. Όταν διαγράφονται με επιτυχία γραμμές από τον πίνακα booking (π.χ. διαγράφονται όλες οι εγγραφές με ημερομηνία μικρότερη του 1970-01-01 00:00:00), τότε οι διαγραμμένες γραμμές εισάγονται αυτόματα στον πίνακα booking-log. Όταν πραγματοποιείται μια ενημέρωση μιας τιμής στον κύριο πίνακα booking, τότε ο trigger ενεργοποιείται και καταγράφει την παλιά εγγραφή στον πίνακα booking-log. Ιδανικά μαζί με τη μεταβαλλόμενη εγγραφή καταγράφει και την ημερομηνία και ώρα της αλλαγής καθώς και το είδος της μεταβολής (d-για delete και u-για update).
- b. Βρείτε και εμφανίστε τα στοιχεία ημερομηνία αναχώρησης, ημερομηνία άφιξης, όνομα αεροδρομίου, κωδικός πτήσης, όνομα επιβάτη, ομαδοποιημένα ανά όνομα επιβάτη και ημερομηνία αναχώρησης. Χρησιμοποιείστε cursors ώστε να εμφανίσετε τις γραμμές σε ομάδες των 15.

Ερώτημα 4 (20%). Σύνδεση ΒΔ με Application Programming Interface (API)

Υλοποιήστε προγραμματιστικά έναν client σε οποιαδήποτε γλώσσα προγραμματισμού γνωρίζετε (π.χ. Python, Java, C) χρησιμοποιώντας την κατάλληλη βιβλιοθήκη σύνδεσης με την PostgreSQL (π.χ. psycopg2, JDBC, ODBC). Ο client θα συνδέεται στο ΣΔΒΔ της PostgreSQL, θα εκτελεί τα queries του Ερωτήματος 2, και θα εμφανίζει τα αποτελέσματα στον χρήστη (είτε σε terminal είτε με γραφικά).

Τρόπος, τόπος και χρόνος παράδοσης – Απορίες σχετικά με την εργασία

Αναρτήστε τις απαντήσεις σας στον παρακάτω σύνδεσμο. Να είστε προσεκτικοί, μόνο μια απάντηση ανά ομάδα (δηλαδή, ένας από κάθε ομάδα θα αναρτήσει την εργασία της ομάδας).

https://docs.google.com/forms/d/e/1FAIpQLScr22EZxKI7OxVx367LoJBvVVPoZR2xkzvkeM6tO2ZpjYwjIQ/viewform?usp=pp_url

Ο σύνδεσμος θα είναι διαθέσιμος μέχρι τις 30/6/2022 23:59. Απαραίτητη διευκρίνιση: **εργασίες δεν γίνονται δεκτές κατά την εξεταστική Σεπτεμβρίου.**

Για οποιαδήποτε απορία αφορά στην εργασία, μπορείτε να απευθυνθείτε στην κα Ρόζα Μαυροπόδη και στον κ. Γιάννη Κοντούλη.

Ζητήματα δεοντολογίας

Είναι προφανές ότι η βαθμολογία πρέπει να αντικατοπτρίζει το επίπεδο της γνώσης που αποκόμισε ο φοιτητής μέσα από το μάθημα και κατάφερε να μεταφέρει αυτή τη γνώση στην εργασία. Για να εξασφαλιστεί όσο είναι δυνατό η παραπάνω αρχή, (α) σε περίπτωση αντιγραφής οι εμπλεκόμενες εργασίες μηδενίζονται, (β) σε περίπτωση αμφιβολίας για το κατά πόσο η ομάδα που αναγράφεται ήταν εκείνη που ανέπτυξε την εργασία, θα κληθεί να την παρουσιάσει για τυχόν διευκρινίσεις. Σημειώνεται ότι το πανεπιστήμιο διαθέτει το λογισμικό λογοκλοπής Turnitin.

Καλή Επιτυχία!

Ακολουθούν ενδεικτικοί πίνακες με αληθοφανή δεδομένα. Προσοχή: **οι παρακάτω πίνακες δεν αντιστοιχούν κατ' ανάγκη στις προδιαγραφές που δόθηκαν στην εκφώνηση, άρα δεν πρέπει να ληφθούν ως έμμεση υπόδειξη για τους πίνακες που πρέπει να υλοποιήσετε.**

```
demo=# select * from bookings limit 5;
```

book_ref	book_date	total_amount
00000F	2017-07-05 03:12:00+03	265700.00
000012	2017-07-14 09:02:00+03	37900.00
000068	2017-08-15 14:27:00+03	18100.00
000181	2017-08-10 13:28:00+03	131800.00
0002D8	2017-08-07 21:40:00+03	23600.00

(5 rows)

```
demo=# select * from aircrafts_data limit 5;
```

aircraft_code	model	range
773	Boeing 777-300	11100
763	Boeing 767-300	7900
SU9	Sukhoi Superjet-100	3000
320	Airbus A320-200	5700
321	Airbus A321-200	5600

(5 rows)

```
demo=# select * from boarding_passes limit 5;
```

ticket_no	flight_id	boarding_no	seat_no
0005435212351	30625	1	2D
0005435212386	30625	2	3G
0005435212381	30625	3	4H
0005432211370	30625	4	5D
0005435212357	30625	5	11A

(5 rows)

```
demo=# select * from flights limit 5;
```

flight_id	flight_no	scheduled_departure	scheduled_arrival	departure_airport	arrival_airport	status	aircraft_code	actual_departure	actual_arrival
1185	PG0134	2017-09-10 09:50:00+03	2017-09-10 14:55:00+03	DME	BTK	Scheduled	319		
3979	PG0052	2017-08-25 14:50:00+03	2017-08-25 17:35:00+03	VKO	HMA	Scheduled	CR2		
4739	PG0561	2017-09-05 12:30:00+03	2017-09-05 14:15:00+03	VKO	AER	Scheduled	763		
5502	PG0529	2017-09-12 09:50:00+03	2017-09-12 11:20:00+03	SVO	UFA	Scheduled	763		
6938	PG0461	2017-09-04 12:25:00+03	2017-09-04 13:20:00+03	SVO	ULV	Scheduled	SU9		

(5 rows)

```
demo=# select * from seats limit 5;
```

aircraft_code	seat_no	fare_conditions
319	2A	Business
319	2C	Business
319	2D	Business
319	2F	Business
319	3A	Business

(5 rows)

```
demo=# select * from ticket_flights limit 5;
```

ticket_no	flight_id	fare_conditions	amount
0005432159776	30625	Business	42100.00
0005435212351	30625	Business	42100.00
0005435212386	30625	Business	42100.00
0005435212381	30625	Business	42100.00
0005432211370	30625	Business	42100.00

(5 rows)

```
demo=# select * from tickets limit 5;
```

ticket_no	book_ref	passenger_id	passenger_name	contact_data
0005432000987	06B046	8149 604011	Apostolou Dimitrios	+30 210 4142476
0005432000988	06B046	8499 420203	Despotis Dimitrios	+30 210 4142315
0005432000989	E170C3	1011 752484	Douligeris Christos	+30 210 4142137
0005432000990	E170C3	4849 400049	Konstantopoulos Ch.	+30 210 4142124
0005432000991	F313DD	6615 976589	Metaxiotis Konstan.	+30 210 4142578

(5 rows)

```
demo=# select * from airports_data limit 5;
```

airport_code	airport_name	city	coordinates	timezone
AAL	Aalborg	Denmark	90001	Europe/...
AES	Aalesund	Norway	90002	Europe/...
AAR	Aarhus	Denmark	90003	Europe/...
YXX	Abbotsford, BC	Canada	90004	USA/...
YXX	Abbotsford, BC	Canada	90005	USA/...

(5 rows)