

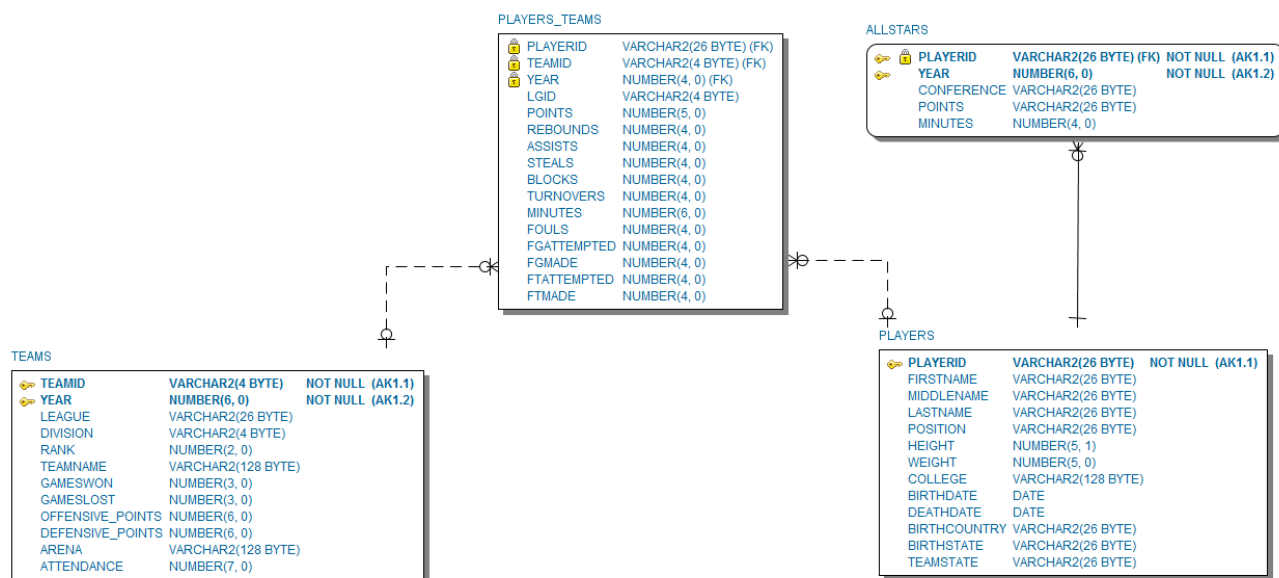
2^η εργασία – PL/SQL - XML

Στόχος των εργασιών είναι η εξοικείωση με θεωρητικά και πρακτικά προβλήματα των Βάσεων Δεδομένων, μέσα από χρηστικά παραδείγματα. Στη δεύτερη εργασία θα ασχοληθούμε με τον προγραμματισμό σε βάσεις δεδομένων και τη γλώσσα PL/SQL για προγραμματισμό σε Oracle. Συγκεκριμένα στόχος της εργασίας είναι η εξοικείωση με τη δημιουργία συναρτήσεων, μεθόδων αλλά και τύπων που ορίζονται από το χρήστη.

A – Ερωτήματα

ΕΡΩΤΗΜΑ 1: Σχήμα εργασίας

Το σχήμα στο οποίο καλείστε να εργαστείτε υπάρχει και πάλι στη ΒΔ xsales και πρέπει να το αντιγράψετε στη βάση σας για να δουλέψετε.



Αφορά ομάδες του αμερικανικού πρωταθλήματος μπάσκετ (teams) και τα συνολικά στατιστικά τους για κάθε χρονιά, παίκτες (players) που έπαιξαν σε αυτές (players_teams) και τα στατιστικά που είχαν κάθε χρονιά καθώς και τις εμφανίσεις κάποιων παικτών στον ετήσιο αγώνα επίδειξης (allstars) και τα στατιστικά τους. Στο σχήμα αποτυπώνονται τα κλειδιά κάθε σχέσης και τα ξένα κλειδιά.

A) Διόρθωση δεδομένων

i) Κατά την αντιγραφή των δεδομένων από την xsales στη βάση σας χάνονται τα constraints (primary και foreign key). Δώστε τις εντολές δημιουργίας τους.

ii) Το ύψος και το βάρος των παικτών στον πίνακα Players έχει υπολογιστεί σε inches και pounds (lb) αντίστοιχα.

Δημιουργήστε μια **procedure** με όνομα **fix_player_metrics** η οποία να μετατρέπει τα inch και pounds σε cm και Kgr αντίστοιχα (1 pound = 0.45 kilograms, 1 inch = 2.54 centimeters).

iii) Οι ομάδες χωρίζονται σε κατηγορίες ανάλογα με τη γεωγραφική τους κατανομή. Η πληροφορία υπάρχει στη στήλη division του πίνακα Teams.

Δημιουργήστε μια **procedure** με όνομα **fix_team_divisions** η οποία θα κάνει τις ακόλουθες αντικαταστάσεις τιμών: AT,CD,SE -->East και SW,PC,NW -->West. Όλες οι υπόλοιπες τιμές να μείνουν ως έχουν.

iv) Ο πίνακας players_teams έχει σύνθετο πρωτεύον κλειδί το playerid, teamid, year. Όπως έχουν μεταφερθεί τα δεδομένα στη βάση υπάρχουν διπλότυπες εγγραφές που δεν επιτρέπουν να δημιουργηθεί κλειδί. Οι διπλές εγγραφές αφορούν παίκτες οι οποίοι έπαιξαν 2 περιόδους την ίδια χρονιά στην ίδια ομάδα συγκεντρώνοντας διαφορετικά στατιστικά (δείτε π.χ. την εγγραφή με playerid='brezepr01').

Δημιουργήστε μια **procedure** με όνομα **fix_duplicate_player_stats** η οποία θα αναλάβει να συγχωνεύσει τα στατιστικά για όσες «διπλές» περιόδους έχει ένας παίκτης, την ίδια χρονιά στην ίδια ομάδα κάνοντας άθροιση των τιμών.

PLAYERID	TEAMID	YEAR	LIGID	POINTS	REBOUNDS	ASSISTS	STEALS	BLOCKS	TURNOVERS	MINUTES	FOULS	FGATTEMPTED	FGMADE	FTATTEMPTED	FTMADE
brezepr01	CHH	2001	NBA	43	28	6	0	7	6	160	28	29	14	25	15
brezepr01	NOH	2002	NBA	42	23	4	2	4	7	110	16	38	15	20	12
brezepr01	NOH	2003	NBA	28	15	3	0	3	6	72	13	26	12	6	4
brezepr01	CHA	2004	NBA	938	531	86	33	55	106	2276	183	756	387	220	164
brezepr01	CHA	2005	NBA	982	440	45	19	32	85	2165	227	791	409	224	164
brezepr01	CHA	2006	NBA	288	184	23	11	23	38	838	161	274	122	68	43
brezepr01	CHA	2007	NBA	37	44	6	0	4	15	267	35	43	17	5	3
brezepr01	CHA	2007	NBA	27	19	4	2	2	5	98	23	13	10	14	7
brezepr01	CHA	2007	NBA	48	18	1	1	2	4	110	25	38	17	21	14
brezepr01	CHA	2009	NBA	5	12	0	1	1	0	36	9	13	2	2	1
brezepr01	CHA	2009	NBA	14	13	1	0	2	2	59	10	13	7	4	0

Συγκεκριμένα η procedure θα κάνει τα ακόλουθα με τη σειρά:

- Θα εντοπίζει σε έναν cursor τις πλειάδες που έχουν διπλότυπα και πρέπει να συγχωνευτούν
- Θα αντιγράφει τις πλειάδες που πρέπει να συγχωνευτούν σε έναν πίνακα duplicate_player_stats (με ίδια δομή με τον players_teams)
- Θα διαγράφει τις εγγραφές που συγχωνεύτηκαν από τον players_teams
- Θα ενημερώνει τον πίνακα players_teams εισάγοντας την νέα συγχωνευμένη εγγραφή

v) Να ορίσετε μια **function** με το όνομα **calc_tendex** η οποία να δέχεται τις κατάλληλες παραμέτρους και να επιστρέφει το δείκτη tendex (πραγματικός αριθμός) για κάποιο παίκτη. Ο δείκτης υπολογίζεται από τον ακόλουθο τύπο:

$$(\text{POINTS} + \text{REBOUNDS} + \text{ASSISTS} + \text{STEALS} + \text{BLOCKS} - \text{Missed FG} - (\text{Missed FT})/2 - \text{TURNOVERS} - \text{FOULS}) / \text{MINUTES}$$

Όπου MissedFG=FGAttempted-FGMade και MissedFT=FTAttempted-FTMade

ΕΡΩΤΗΜΑ 2: Εξαγωγή σε XML

i) Να ορίσετε μια **procedure** με όνομα **get_top_players_xml** που να δέχεται ως όρισμα το έτος και έναν αριθμό N και να παράγει ένα XML με τους N παίκτες με το μεγαλύτερο tendex index για το συγκεκριμένο έτος σε Ανατολή και τους N καλύτερους σε Δύση. Να εμφανίζει playerid, όνομα, επώνυμο, position, πόντους, λεπτά συμμετοχής, το δείκτη index, division και την ομάδα στην οποία έπαιξαν το έτος αυτό (όνομα και teamid).

Καλέστε την procedure αυτή για N=12 και Year=2009 και αποθηκεύστε την έξοδο στο αρχείο topplayers.xml

ii) Να ορίσετε μια **procedure** όνομα **get_allstar_players_xml** που να επιλέγει τους allstar παίκτες του 2009 και να επιλέγει γι' αυτούς την ακόλουθη πληροφορία (conference=division):

CONFERENCE	POINTS	MINUTES	FIRSTNAME	LASTNAME	POSITION	TEAMNAME
West	27	29	Carmelo	Anthony	F	Denver Nuggets
West	17	26	Chauncey	Billups	G	Denver Nuggets
East	23	29	Christopher	Bosh	F	Toronto Raptors
West	3	13	Timothy	Duncan	C-F	San Antonio Spurs
West	15	20	Kevin	Durant	F	Oklahoma City Thunder
East	4	13	Kevin	Garnett	F	Boston Celtics
West	13	20	Pau	Gasol Saez	F	Los Angeles Lakers
East	8	14	Alfred	Horford Reynoso	F-C	Atlanta Hawks
East	17	27	Dwight	Howard	F	Orlando Magic
East	25	32	LeBron	James	F	Cleveland Cavaliers
East	10	18	Joe	Johnson	G-F	Atlanta Hawks
West	4	11	Christopher	Kaman	C	Los Angeles Clippers
West	0	6	Jason	Kidd	G	Dallas Mavericks
East	4	12	David	Lee	F	New York Knicks
West	4	20	Stephen	Nash	G	Phoenix Suns
West	22	28	Dirk	Nowitzki	C	Dallas Mavericks
East	8	11	Paul	Pierce	F	Boston Celtics
West	8	19	Zachary	Randolph	F	Memphis Grizzlies
East	4	20	Rajon	Rondo	F	Boston Celtics
East	8	15	Derrick	Rose	G	Chicago Bulls
West	12	20	Amar'e	Stoudemire	F-C	Phoenix Suns
East	28	31	Dwyane	Wade	G	Miami Heat
East	2	16	Gerald	Wallace	F	Charlotte Bobcats
West	14	28	Deron	Williams	G	Utah Jazz

Να παράγει ένα xml αρχείο παρόμοιο με το προηγούμενο στο οποίο φυσικά θα λείπει η πληροφορία του tendex.

Τα αρχεία XML θα έχουν την ακόλουθη μορφή:

```

▼<nba dataset="allstars"> ▼<nba dataset="topplayers">
  ▼<division id="east"> ▼<division id="east">
    <player></player> <player></player>
    <player></player> <player></player>
  </division> </division>
  ▼<division id="west"> ▼<division id="west">
    <player></player> <player></player>
    <player></player> <player></player>
  </division> </division>
</nba> </nba>

```

Επιλέξτε εσείς τη δομή εντός του element player.

lii) Να δώσετε το αρχείο DTD που περιγράφει τη δομή των δύο αρχείων σας.

ΕΡΩΤΗΜΑ 3: Ερωτήματα με χρήση XPath

Συγχωνεύστε τα δύο xml αρχεία σε ένα (merged.xml), βάζοντάς τα κάτω από μια κοινή ρίζα. Στη συνέχεια να απαντήσετε στα ακόλουθα ερωτήματα με χρήση XPath. Σε κάθε ερώτημα εκτός από την εντολή XPath να δώσετε και το τελικό αποτέλεσμα.

- 1) Εμφανίστε όλους τους παίκτες της Ανατολής που πέτυχαν στο allstar παιχνίδι πάνω από 10 πόντους.
- 2) Εμφανίστε όλους τους παίκτες που έχουν πάνω από 0.7 tendex σε ανατολή και δύση.
- 3) Εμφανίστε τις ομάδες από τις οποίες προέρχονται οι allstar παίκτες της Δύσης.

Για τις δοκιμές σας μπορείτε να χρησιμοποιήσετε ένα από τα:

- XML Copy Editor <http://xml-copy-editor.sourceforge.net/>
- Online XPath Query Checker: <http://www.xpathtester.com/xpath>
- XPath Expression Testbed:
<http://www.whitebeam.org/library/guide/TechNotes/xpathtestbed.rhtm>

ή όποιο άλλο νομίζετε εσείς

B – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των **3 ατόμων (το πολύ)**, αν και επιτρέπονται μικρότερες ομάδες. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους ΑΜ των μελών της ομάδας: π.χ. **AM1-AM2.zip**

- Το zip θα περιλαμβάνει:
 - ένα αρχείο readme.txt
 - με τα ονοματεπώνυμα και τους ΑΜ των φοιτητών της ομάδας
 - όσα επιπλέον αρχεία δημιουργήσετε (sql scripts, xml, dtd)
 - το αρχείο word με την τελική εργασία

Γ – Άλλες Οδηγίες

Όσες εργασίες δεν τηρούν τις οδηγίες παράδοσης, θα έχουν επίπτωση στο βαθμό.

Όσες εργασίες κριθούν ότι είναι **αντιγραφές θα μηδενίζονται**.

Ημερομηνία παράδοσης: **Στο e-class με οριστική τελική ημερομηνία 13-1-2017**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 15-1-2017 θα έχουν μείωση 2 μονάδων στο βαθμό.