Ανάλυση Δεδομένων με χρήση εργαλείων ΒΙ

Δημήτρης Δασκόπουλος, ΚΗΔ ΑΠΘ



Σκοπιμότητα

- Πρόσβαση σε πλήρη δεδομένα του SIS
- Ανάλυσης των δεδομένων καθημερινής λειτουργίας του ιδρύματος
- Προσφέρει μετρικές για αθέατες παραμέτρους

Ανάλυση Δεδομένων

Συνήθη εμπόδια

- Κλειστή πρόσβαση στα δεδομένα
- Περιορισμένη γνώση υποκείμενων δεδομένων, μοντέλα χωρίς οργάνωση
- Κλειστές εφαρμογές σχεδιασμένες από άλλους
- Παρουσίαση στατικών αριθμών και πινάκων συνήθως σε μορφή αναφορών

Εργαλεία ΒΙ

- Πρότυπα πρωτόκολλα πρόσβασης σε πηγές δεδομένων
- Εύκολη εξερεύνηση υποκείμενων δεδομένων
- Δυνατότητα δημιουργίας custom dashboards σε όποια μορφή επιθυμεί ο χρήστης
- Εξαιρετική οπτικοποίηση με διαγράμματα
- Διαδραστική άντληση δεδομένων με φίλτρα



Άντληση Δεδομένων στο UniverSIS

- ▶ ΒΔ με οργανωμένα και τυποποιημένα ονόματα μοντέλων και πεδίων
- Εφαρμογές έχουν πρόσβαση μόνο μέσω του API δεδομένα βάσει ρόλου χρήστη
- Δυναμικές μέθοδοι API (data models, query language, εκτενής τεκμηρίωση)
- ► Object Relational Mapping (ORM) και Object Oriented Programming (OOP)
- Απουσία entity relations σε επίπεδο SQL (λόγω ORM και OOP)
- Εξετάζουμε δύο εναλλακτικές
 - Database
 - API



Σύγκριση μεθόδων άντλησης δεδομένων

Database

- +
- Σχέση μεταξύ μοντέλων
- Φιλικά ονόματα μοντέλων
- Μηχανισμοί ανωνυμοποίησης
- -
- Πρόσβαση σε δεδομένα παρακάμπτοντας μηχανισμούς ασφαλείας
- Δυσκολία διαμοιρασμού δεδομένων σε τρίτους

API

- -
 - Πρόσβαση σε δεδομένα βάσει του ρόλου σου
 - Διαμοιρασμός δεδομένων βάσει αρμοδιότητας
 - Ευέλικτο πρωτόκολλο ερωτημάτων (ODATA)
- -
- Δυσκολία διασύνδεσης μοντέλων
- Ανάγκη μετονομασίας μοντέλων
- Επεξεργασία JSON data



Ανωνυμοποίηση Δεδομένων

Εναλλακτικές μέθοδοι για περιορισμό SQL user του PowerBI

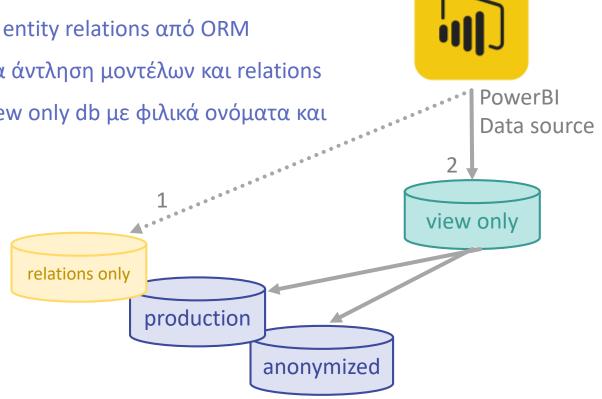
- MSSQL Dynamic Masking σε πρωτογενή ΒΔ
 - αποκλεισμός πρόσβασης σε πεδία για κάποιους χρήστες
- Τροποποίηση δεδομένων σε αντίγραφο BΔ
 - τροποποίηση δεδομένων με πακέτο
- ▶ DB replication για ζωντανό αντίγραφο BΔ
 - Περιορισμός πρόσβασης με permission rules



Database ως πηγή για PowerBI

Βήματα

- UniverSIS: Αυτόματη δημιουργία κενής db για entity relations από ORM
- PowerBI: Ορισμός data source από κενή db για άντληση μοντέλων και relations
- PowerBI: Αλλαγή data source σε ενδιάμεση view only db με φιλικά ονόματα και εναλλακτικές πηγές
 - production db
 - anonymized db





API ως πηγή για PowerBI

Βήματα

- UniverSIS: Απόκτηση OAUTH token μετά από login
- PowerBI: Δημιουργία data connector μέσω UniverSIS API
- PowerBI: Επεξεργασία των JSON data για επιλογή πεδίων

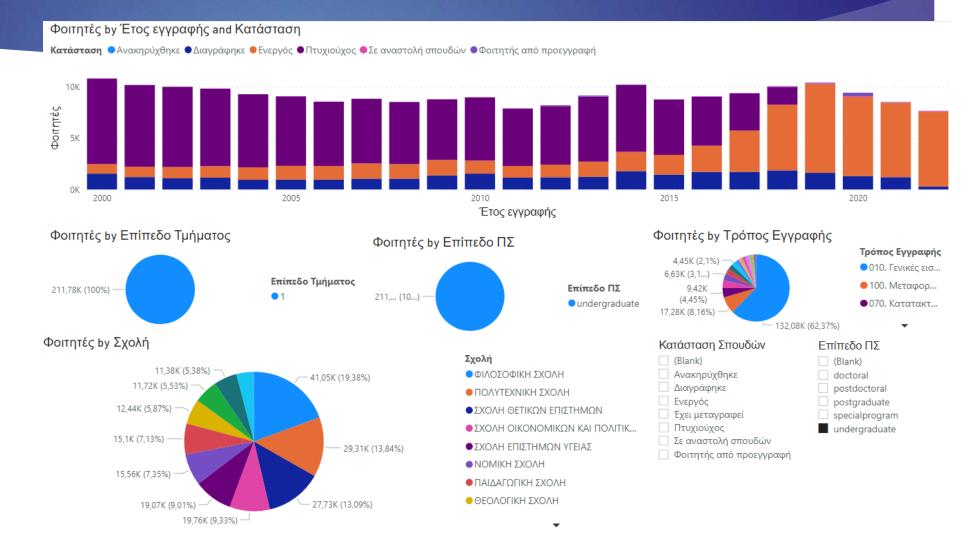
Εναλλακτικά άντληση πινάκων δεδομένων από queries προς τον ΑΡΙ με τη μορφή .xlsx ή .csv



Παραδείγματα Visualization Dashboards

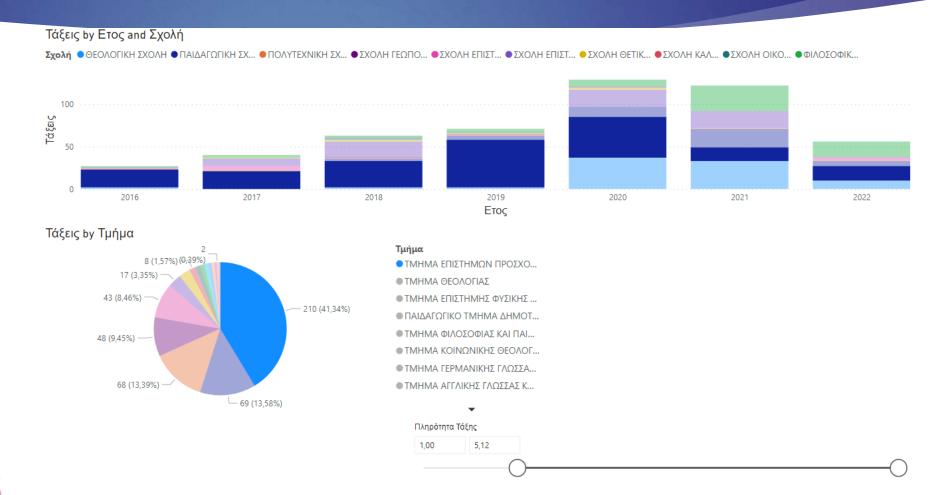
- Κατάσταση φοιτητών ανά έτος εισαγωγής
- Εισακτέοι ανά τρόπο εγγραφής
- Πληρότητα σε τάξεις με μέγιστο όριο
- Μέσος αριθμός περασμένων ECTS ανά έτος
- Μέσος όρος βαθμού πτυχίου ανά τμήμα
- Κατανομές βαθμών ανά Τμήμα
- Χρόνος εξυπηρέτησης αιτήσεων πιστοποιητικών







Πληρότητα τάξεων με όριο θέσεων





Ερωτήσεις

Ευχαριστούμε