**Nazwa przedmiotu:**

Projektowanie hurtowni danych

**Koordynator przedmiotu:**

dr inż. Katarzyna Rostek

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia I stopnia

**Program:**

Zarządzanie i Inżynieria Produkcji

**Grupa przedmiotów:**

Technologie informatyczne

**Kod przedmiotu:**

PROHU

**Semestr nominalny:**

4 / rok ak. 2010/2011

**Liczba punktów ECTS:**

2

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

**Język prowadzenia zajęć:**

polski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład:  | 15h |
| Ćwiczenia:  | 0h |
| Laboratorium:  | 15h |
| Projekt:  | 0h |
| Lekcje komputerowe:  | 0h |

**Wymagania wstępne:**

Słowa kluczowe:
bazy danych, hurtownie danych, składnice danych, ROLAP, MOLAP, OLAP, HOLAP, analiza OLAP, eksploracja danych, fakty, wymiary, agregacje, metadane.

**Limit liczby studentów:**

**Cel przedmiotu:**

Przedmiot składa się z jednostki wykładowej oraz laboratorium. Wykłady oparte są na prezentacjach multimedialnych prezentowanych przez prowadzącego. Elementem pracy twórczej w ramach wykładu jest tworzony w zespołach 2-osobowych projekt hurtowni danych dla wybranego problemu decyzyjnego oraz wybranego środowiska projektowego. Laboratorium oparte jest na ćwiczeniach wykonywanych indywidualnie, w ramach których tworzona jest od podstaw składnica danych i realizowany jest pakiet analiz OLAP oraz eksploracji danych. Elementem twórczego wkładu studentów jest przygotowanie raportu końcowego z samodzielnie wydedukowanymi i opracowanymi wnioskami z wykonanych analiz.

**Treści kształcenia:**

Wykład 15h:
1h - Wprowadzenie do zagadnień hurtowni danych:
Geneza powstania hurtowni danych. Porównanie hurtowni danych i baz transakcyjnych. Definicja hurtowni danych. Cechy charakterystyczne hurtowni danych.
1h - Przykłady rzeczywistych wdrożeń hurtowni danych:
Racjonalne przesłanki wdrożenia hurtowni danych w przedsiębiorstwie. Korzyści wdrożenia hurtowni danych w przedsiębiorstwie na wybranych przykładach firm polskich i zagranicznych.
1h - Typy struktur modelu danych hurtowni:
Typy danych hurtowni: fakty, wymiary, agregacje, metadane. Typy struktur danych: Rolap, Molap, Holap.
1h - Projektowanie struktury modelu danych hurtowni:
Metodyka projektowania struktur tabel faktów i wymiarów w modelu Rolap.
1h - Typy architektury hurtowni danych :
Czynniki determinujące dobór architektury hurtowni danych w przedsiębiorstwie. Architektura niezależnych składnic danych. Architektura korporacyjnej hurtowni danych. Architektura CIF Bill’a Inmon’a. Architektura BUS Ralpha Kimball’a. Architektura przyrostowa. Architektura federacyjna.
1h - Różne podejścia do projektowania architektury hurtowni danych:
Fazy projektowania hurtowni danych. Metodyka „od ogółu do szczegółu”. Metodyka „od szczegółu do ogółu”. Metodyka federacyjna.
1h - Korzyści i zagrożenia wdrożenia hurtowni danych:
Ryzyko inwestycyjne. Ryzyko organizacyjne i biznesowe. Ryzyko technologiczne. Cechy efektywnego i użytecznego projektu hurtowni danych.
1h - Metodyka wdrożenia hurtowni danych:
Podejście całościowe vs. podejście iteracyjne.
1h - Systemy informatyczne z hurtownią danych:
Systemy analityczno-raportujące. Systemy Business Intelligence. Systemy Corporate Performance Management. Systemy wspomagania decyzji.
1h - Wprowadzenie do zagadnień przetwarzania analitycznego danych:
Typy przetwarzania analitycznego danych. Języki przetwarzania analitycznego danych. Technologie analitycznego przetwarzania danych.
1h - Przetwarzanie analityczne na bieżąco – OLAP:
Definicja OLAP E. Codd’a. 12 cech systemów OLAP według Codd’a. Definicja OLAP FASMI. Zakres zastosowań systemów OLAP. Przykłady rzeczywistych wdrożeń systemów OLAP.
1h - Eksploracja danych - data mining:
Definicja eksploracji danych. Podział klasyfikacyjny analiz eksploracji danych. Metody i techniki eksploracji danych. Zakres zastosowań systemów eksploracji danych. Przykłady rzeczywistych wdrożeń systemów eksploracji danych.
1h - Metodyka realizacji procesu pozyskiwania wiedzy z danych:
Definicja procesu pozyskiwania wiedzy z danych. 13 etapowa metodyka realizacji procesu pozyskiwania wiedzy z danych.
1h - Przykłady zastosowań procesu pozyskiwania wiedzy z danych:
Zakres zastosowań procesu pozyskiwania wiedzy z danych. Przykład zastosowania procesu pozyskiwania wiedzy w firmie branży ubezpieczeniowej.
1h - Test zaliczeniowy:
Test 10 otwartych pytań: 5 pytań problemowych oraz 5 pytań teoretycznych.

Laboratorium 15h:
1h - Poznanie i przygotowanie zbiorów źródłowych danych pod kątem projektu hurtowni danych:
Przygotowanie 3 różnych źródeł operacyjnych pod kątem ich wspólnej integracji w modelu danych hurtowni. Sprawdzenie jakości danych źródłowych.
1h - Implementacja struktur hurtowni:
Implementacja hurtowni danych w modelu gwiazdy.
1h - Ekstrakcja danych ze zbiorów źródłowych:
Tworzenie połączenia ekstrakcyjnego pomiędzy danymi źródłowymi a bazą danych hurtowni.
1h - Transformacja i ładowanie danych do struktur hurtowni:
Przekształcenie danych do postaci właściwej dla modelu danych hurtowni. Ładowanie danych do modelu danych hurtowni.
1h - Implementacja i generowanie kostek OLAP:
Definiowanie struktur kostek wielowymiarowych OLAP. Generowanie danych w kostkach wielowymiarowych OLAP.
1h - Projekt analizy przetwarzania analitycznego danych na bieżąco:
Przygotowanie pakietu analiz OLAP.
1h - Przetwarzanie analityczne danych na bieżąco:
Realizacja analiz OLAP według przygotowanego projektu.
1h - Analiza wyników przetwarzania analitycznego danych na bieżąco:
Analiza otrzymanych wyników i zamieszczenie ich w raporcie zaliczeniowym.
1h - Projekt analizy eksploracji danych:
Zaprojektowanie modelu eksploracji danych.
1h - Implementacja struktur tablicy eksploracji danych:
Przygotowanie projektu zbioru źródłowego dla modelu eksploracji danych.
1h - Ekstrakcja, transformacja i ładowanie danych ze zbiorów źródłowych do tablicy eksploracji danych:
Generowanie zbioru źródłowego dla modelu eksploracji danych.
1h - Eksploracja danych:
Implementacja modelu eksploracji danych.
1h - Analiza wyników eksploracji danych:
Ewaluacja modelu. Poszukiwanie modelu suboptymalnego.
1h - Opracowanie wyników projektu laboratoryjnego:
Opracowanie wyników wszystkich zrealizowanych ćwiczeń laboratoryjnych w raporcie końcowym.
1h - Rozliczenie wyników projektu laboratoryjnego:
Prezentacja wyników końcowych zrealizowanych podczas ćwiczeń laboratoryjnych.

**Metody oceny:**

Wykład – projekt wykonywany w grupach 2-osobowych oraz test, składający się z 10 otwartych pytań, sprawdzający wiedzę teoretyczną.
Laboratorium – projekt wykonywany w grupach 2-osobowych, rozliczany na podstawie przygotowanego raportu końcowego.

**Egzamin:**

**Literatura:**

Literatura obowiązkowa:
• Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P. - Hurtownie danych. Podstawa organizacji i funkcjonowania. WSiP, Warszawa, 2003.
• Poe V., Klauer P., Brobst S. - Tworzenie hurtowni danych. WNT, Warszawa, 2000.
• Todman Ch. - Projektowanie hurtowni danych. WNT, Warszawa, 2003.

Literatura zalecana:
• Inmon W. H. - Building the Data Warehouse. 4th Edition, John Wiley & Sons, 2005.
• Kimball R., Ross M., Thornthwaite W., Mundy J., Becker B. - The Data Warehouse Lifecycle Toolkit. 2nd Edition: Practical Techniques for Building Data Warehouse and Business Intelligence Systems. John Wiley & Sons, 2008.
• Inmon W.H., Strauss D., Neushloss G. - DW 2.0: The Architecture for the Next Generation of Data Warehousing. Morgan Kaufmann, 2008.

Oprogramowanie:
• Sybase Power Designer,
oraz do wyboru:
• SAS Data Integration Studio, SAS OLAP Cube Studio, SAS Enterprise Guide, SAS Enterprise Miner,
• SYBASE IQ,
• MS SQL Server Enterprise Edition.

**Witryna www przedmiotu:**

**Uwagi:**

## Efekty przedmiotowe