



## KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Hurtownie danych [N2Inf1-IWPB>HD]

### Przedmiot

Kierunek studiów  
Informatyka

Rok/Semestr  
1/2

Studia w zakresie (specjalność)  
Informatyka w procesach biznesowych

Profil studiów  
ogólnoakademicki

Poziom studiów  
drugiego stopnia

Język oferowanego przedmiotu  
polski

Forma studiów  
niestacjonarne

Wymagalność  
obieralny

### Liczba godzin

Wykład  
16

Laboratorium  
16

Inne  
0

Ćwiczenia  
0

Projekty/seminaria  
0

### Liczba punktów ECTS

3,00

### Koordynatorzy

dr hab. Tadeusz Pankowski  
tadeusz.pankowski@put.poznan.pl

### Wykładowcy

### Wymagania wstępne

Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z systemów baz danych i języków programowania. Powinien posiadać umiejętność rozwiązywania podstawowych problemów z projektowania systemów informatycznych, administrowania systemami baz danych, formułowania poleceń w języku SQL oraz umiejętność pozyskiwania informacji ze wskazanych źródeł. Powinien również rozumieć konieczność poszerzania swoich kompetencji / mieć gotowość do podjęcia współpracy zespołowej. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.

## Cel przedmiotu

Wskazanie praktycznych problemów jakie rozwiązuje się projektując, implementując i wdrażając systemy hurtowni danych oraz analizy danych biznesowych (BI). Przekazanie studentom wiedzy dotyczącej projektowania systemów hurtowni danych, w zakresie: architektury, modelowania danych, projektowania warstwy integrującej i zasilającej - ETL, struktur fizycznych, zarządzania metadanymi, trendów rozwojowych hurtowni danych i systemów BI. Kształcenie umiejętności w zakresie formułowania problemów analizy danych oraz posługiwania się językami i narzędziami wspomagającymi analizę danych. Przedstawienie problematyki implementowania hurtowni danych i aplikacji klasy BI. Rozwijanie u studentów umiejętności rozwiązywania problemów, w zakresie: projektowania i implementowania systemu hurtowni danych, oceny przydatności technologii HD i BI do konkretnego zastosowania. Kształtowanie u studentów umiejętności pracy zespołowej w ramach budowy systemów hurtowni danych i systemów analizy danych biznesowych. Kształtowanie u studentów umiejętności realizowania projektów HD i BI. Kształtowanie u studentów umiejętności korzystania ze źródeł wiedzy (np. publikacje anglojęzyczne) i samokształcenia się oraz zarządzania czasem.

## Przedmiotowe efekty uczenia się

Wiedza:

1. ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie architektury systemów informatycznych, w szczególności hurtowni danych i baz danych (K2st\_W4)
2. ma podbudowaną teoretycznie szczegółową wiedzę związaną z wybranymi zagadnieniami z zakresu informatyki, takimi jak: architektury i technologie integracji danych, metodyka projektowania systemu informatycznego (w szczególności hurtowni danych), technologie analizy danych, modele składowania danych (K2s-W5)
3. ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w dziedzinie przetwarzania danych i hurtowni danych (K2s-W6)
4. ma podstawową wiedzę o cyklu życia systemów informatycznych, w szczególności systemów hurtowni danych (K2s-W7)
5. zna podstawowe metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z wybranego obszaru informatyki (projektowania i implementowania systemu hurtowni danych, projektowania i implementowania schematów hurtowni danych, analizy danych, wizualizacji raportów, projektowania zapytań analitycznych) (K2s-W8)

Umiejętności:

W wyniku przeprowadzonych zajęć student powinien wykazać się umiejętnościami w zakresie (student potrafi):

1. pozyskiwać informacje z literatury, baz danych oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie (także przygotowywać prezentacje i raporty techniczne) (K2s-U1)
2. określić kierunki dalszego uczenia się i zrealizować proces samokształcenia (K2s-U5)
3. wykorzystać do formułowania i rozwiązywania zadań inżynierskich i prostych problemów badawczych metody eksperymentalne (K2s-U9)
4. integrować wiedzę z różnych obszarów informatyki (np. inżynieria oprogramowania, administrowanie systemami informatycznymi, bazy danych) oraz zastosować podejście systemowe, uwzględniające także aspekty pozatechniczne (K2s-U10)
5. formułować i testować hipotezy związane z problemami inżynierskimi i prostymi problemami badawczymi (K2s-U12)
6. ocenić przydatność i możliwość wykorzystania nowych osiągnięć (metod i narzędzi) oraz nowych produktów informatycznych (w szczególności systemów zarządzania bazami danych, środowisk programistycznych ETL)(K2s-U13)

Kompetencje społeczne:

W wyniku przeprowadzonych zajęć student zdobędzie wymienione niżej kompetencje. Zaliczenie przedmiotu oznacza, że student:

1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe (K2s-K1)
2. zna przykłady i rozumie przyczyny wadliwie i właściwie działających systemów informatycznych (w szczególności systemów hurtowni danych) (K2s-K4)
3. potrafi odpowiednio określić priorytety służące realizacji zadania określonego przez siebie lub innych (K2s-K6)
4. potrafi współdziałać i pracować w grupie, przyjmując w niej różne role (m.in., kierownika projektu,

analitka, programisty) (K2s-K5)

## Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

a) w zakresie wykładów:

- na podstawie odpowiedzi na pytania dotyczące materiału omówionego na wykładach,

b) w zakresie laboratoriów:

- na podstawie oceny bieżącego postępu realizacji zadań,

Ocena podsumowująca:

a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- ocenę wiedzy i umiejętności wykazanych w formie testu z pytaniami otwartymi oraz pytaniami jedno- lub wielokrotnego wyboru;

Test składa się z szeregu pytań (15-30) otwartych oraz jedno lub wielokrotnego wyboru. Test uznaje się za zaliczony po uzyskaniu ponad 50% możliwych do zdobycia punktów. Przyjmuje się następującą skalę ocen i punktów: <0;50%>: ndst., (50%;60%>: dst, (60%;70%>: dst+, (70%;80%>: db, (80%;90%>: db+, (90%;100%>: bdb

- omówienie wyników testu;

b) w zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:

- ocenę realizacji zadań zleczanych na każdych zajęciach,
- ocenę wiedzy i umiejętności związanych z realizacją zadań laboratoryjnych poprzez rozwiązanie jednego kolokwium na koniec semestru.
- uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:
  - o omówienia dodatkowych aspektów zagadnienia,
  - o uwagi związane z udoskonaleniem materiałów dydaktycznych.

## Treści programowe

Program wykładów obejmuje następujące zagadnienia:

1) wprowadzenie do problematyki analizy danych, hurtowni danych i business intelligence (uzasadnienie potrzeby wykorzystywania tych technologii), 2) systemy transakcyjne OLTP vs. Systemy analityczne OLTP, stos technologiczny SQLBI, 3) architektury systemu hurtowni danych, 4) modelowanie danych (konceptualny model wielowymiarowy, model tabelaryczny, schematy danych: gwiazdy, płotka śniegu, konstelacji faktów), 5) problematyka modelowania wymiarów ze szczególnym uwzględnieniem problematyki wymiarów wolnozmiennych, 6) elementy składowe modelu danych w HD: tabele faktów, tabele wymiarów, miary, wskaźniki KPI, 7) problematyka integracji danych dla hurtowni danych, projektowanie warstwy ETL, 8) projektowanie przepływów danych w procesach ETL, 9) wykorzystanie języka SQL i funkcji analitycznych w analizie danych, 10) model tabelaryczny: znaczenie i wykorzystanie języka DAX (Data Analysis Expressions), 11) tworzenie raportów i pulpitów nawigacyjnych (dashboardów) jako wyników analizy, 12) modele wielowymiarowe: tworzenie kostek i operowanie na kostce w języku MDX (MultiDimensional Expressions), 13) wybrane struktury fizyczne i problemy optymalizacji.

Zajęcia laboratoryjne prowadzone są w formie czterech 4-godzinnych zajęć, odbywających się w laboratorium. Program zajęć podzielono na osiem części: 1. Wprowadzenie do środowiska laboratoryjnego. 2. Tworzenie tabelarycznego modelu danych i jego analiza w środowisku Power BI Desktop. 3. Zaawansowane funkcje analizy danych z wykorzystaniem języka DAX. 4. Tworzenie hurtowni danych w środowisku SQL Server, zastosowanie funkcjonalności SQL w analizie danych. 5. Procesy ETL w środowisku SQL Server Integration Services. 6. Tworzenie tabelarycznego modelu analizy danych dla hurtowni danych ( SQL Server Analysis Services i Visual Studio SSDT w modelowaniu tabelarycznym). 7. Definiowanie miar, kolumn obliczanych i wskaźników KPI, tworzenie raportów i pulpitów nawigacyjnych. 8. Tworzenie modelu wielowymiarowego i kostek danych, zastosowanie języka MDX i Excel Power Pivot. 9. Integrowanie współpracy poszczególnych składowych stosu technologicznego SQLBI. Rozważanie problemów tworzenia raportów i wizualizacji wyników analizy za pomocą pulpitów nawigacyjnych.

Zajęcia są prowadzone w formie zajęć ćwiczeniowych przy komputerach, przy czym każdy student pracuje samodzielnie. Każde zadanie jest poprzedzone krótką prezentacją a następnie omówione zagadnienia są ćwiczone w praktyce.

## Tematyka zajęć

Program wykładów obejmuje następujące zagadnienia:

1) wprowadzenie do problematyki analizy danych, hurtowni danych i business intelligence (uzasadnienie potrzeby wykorzystywania tych technologii), 2) systemy transakcyjne OLTP vs. Systemy analityczne OLTP, stos technologiczny SQLBI, 3) architektury systemu hurtowni danych, 4) modelowanie danych (konceptualny model wielowymiarowy, model tabelaryczny, schematy danych: gwiazdy, płatka śniegu, konstelacji faktów), 5) problematyka modelowania wymiarów ze szczególnym uwzględnieniem problematyki wymiarów wolnozmiennych, 6) elementy składowe modelu danych w HD: tabele faktów, tabele wymiarów, miary, wskaźniki KPI, 7) problematyka integracji danych dla hurtowni danych, projektowanie warstwy ETL, 8) projektowanie przepływów danych w procesach ETL, 9) wykorzystanie języka SQL i funkcji analitycznych w analizie danych, 10) model tabelaryczny: znaczenie i wykorzystanie języka DAX (Data Analysis Expressions), 11) tworzenie raportów i pulpitów nawigacyjnych (dashboardów) jako wyników analizy, 12) modele wielowymiarowe: tworzenie kostek i operowanie na kostce w języku MDX (MultiDimensional Expressions), 13) wybrane struktury fizyczne i problemy optymalizacji.

Zajęcia laboratoryjne prowadzone są w formie czterech 4-godzinnych zajęć, odbywających się w laboratorium. Program zajęć podzielono na osiem części: 1. Wprowadzenie do środowiska laboratoryjnego. 2. Tworzenie tabelarycznego modelu danych i jego analiza w środowisku Power BI Desktop. 3. Zaawansowane funkcje analizy danych z wykorzystaniem języka DAX. 4. Tworzenie hurtowni danych w środowisku SQL Server, zastosowanie funkcjonalności SQL w analizie danych. 5. Procesy ETL w środowisku SQL Server Integration Services. 6. Tworzenie tabelarycznego modelu analizy danych dla hurtowni danych ( SQL Server Analysis Services i Visual Studio SSDT w modelowaniu tabelarycznym). 7. Definiowanie miar, kolumn obliczanych i wskaźników KPI, tworzenie raportów i pulpitów nawigacyjnych. 8. Tworzenie modelu wielowymiarowego i kostek danych, zastosowanie języka MDX i Excel Power Pivot. 9. Integrowanie współpracy poszczególnych składowych stosu technologicznego SQLBI. Rozważanie problemów tworzenia raportów i wizualizacji wyników analizy za pomocą pulpitów nawigacyjnych.

Zajęcia są prowadzone w formie zajęć ćwiczeniowych przy komputerach, przy czym każdy student pracuje samodzielnie. Każde zadanie jest poprzedzone krótką prezentacją a następnie omówione zagadnienia są ćwiczone w praktyce.

## Metody dydaktyczne

1. wykład: prezentacja multimedialna, prezentacja jest uzupełniana krótkimi przykładami prezentowanymi w sposób tradycyjny z wykorzystaniem tablicy, rozwiązywanie problemów i omawianie rozwiązań na tablicy, wiele problemów omawianych jest metodą prezentacji na komputerze, 2. laboratoria: prezentacja multimedialna, prezentacja jest uzupełniana krótkimi przykładami prezentowanymi w sposób tradycyjny z wykorzystaniem tablicy, wykonywanie ćwiczeń w hurtowni danych, omawianie trudniejszych ćwiczeń przy tablicy, odpowiedzi na pytania na bieżąco, rozwiązywanie problemów na bieżąco, dominuje praca na komputerach.

## Literatura

Podstawowa

[1] M., Jarke, M., Lenzerini, Y., Vassiliou, Hurtownie danych. Podstawy organizacji i funkcjonowania , Wydawnictwa Szkolne i Pedagogiczne, Warszawa, 2003.

[2] M. Russo, A. Ferrari, Kompletny przewodnik po DAX. Analiza biznesowa przy użyciu Microsoft Power BI, SQL Server Analysis Services i Excel, APN Promise, Warszawa 2019.

Uzupełniająca

[1] Ch. Todman, Projektowanie hurtowni danych, WNT, Warszawa 2003

[2] R. Kimball, J. Caserta: The Data Warehouse ETL Toolkit. Wiley, 2004

[3] A. Chądzyńska-Krasowska, E. Mrówka-Matejewska, M. Jankowski-Lorek, Podstawy hurtowni danych, Wydawnictwo Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych, Warszawa 2013.

## Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	75	3,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	32	1,50
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwiiw/egzaminu, wykonanie projektu)	43	1,50