

Uczelniana Oferta Studiów Zaawansowanych
SYLABUS 2014/2015

Nazwa przedmiotu	Analiza danych z pakietem R (grupa I – wersja podstawowa)
Liczba punktów ECTS	Proponowana liczba punktów: 3 ECTS, zatwierdza dziekan danego wydziału

Osoby prowadzące	Tytuł naukowy	Imię i nazwisko	Katedra / Instytut / Centrum / Inne
	Dr hab. inż.	Anna Dembińska	Zakład Procesów Stochastycznych i Matematyki Finansowej, Wydział MiNI PW
Osoba odpowiedzialna za przedmiot	Dr hab. inż.	Anna Dembińska	Zakład Procesów Stochastycznych i Matematyki Finansowej, Wydział MiNI PW

Semestr studiów	Semestr letni 2014/2015
Typ przedmiotu (możliwości wyboru) obowiązkowy O fakultatywny F	F
Wymagania wstępne Zakres wiadomości / kompetencji / umiejętności, jakie powinien już posiadać student przed rozpoczęciem nauki przedmiotu, a także specyfikacja innych przedmiotów lub programów, które należy zaliczyć wcześniej. Uwaga: maksymalna objętość tekstu to 1/2 standardowej strony A4	Podstawy z rachunku prawdopodobieństwa: pojęcie prawdopodobieństwa i jego własności, niezależność zdarzeń losowych, podstawowe rozkłady prawdopodobieństwa ciągłe i dyskretne, podstawowe wiadomości dotyczące zmiennych losowych jedno- i wielowymiarowych, Centralne Twierdzenie Graniczne. Podstawy ze statystyki: znajomość podstawowych statystyk próbkowych i podstawowych metod graficznej prezentacji danych, znajomość pojęcia testu statystycznego.
Poziom przedmiotu Podstawowy P Średniozaawansowany Ś Zaawansowany Z	Ś

CSZ sylabus 2014/2015/ 1
Wykład współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego

KAPITAŁ LUDZKI
 NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
 EUROPEJSKI
 FUNDUSZ SPOŁECZNY


Charakter zajęć , liczba godzin w semestrze, liczba godzin w tygodniu. 1) podać rodzaj prowadzonych zajęć dla danego przedmiotu: wykłady (W); ćwiczenia (Ć); laboratorium (L); projekt (P) 2) podać liczbę godzin w tygodniu np. W - 2; Ć - 2; L - 3; P - 0 3) podać liczbę godzin w semestrze np. W - 30; Ć - 30; L - 45; P - 0	L – 10 spotkań po 3 godz., łącznie: 30 godzin
Sugerowana liczba godzin pracy własnej	45 godzin obejmuje : 30 godzin przygotowywanie się słuchacza do zajęć, 15 godzin – analiza i rozwiązywanie projektów
Całkowita liczba godzin:	75 godzin
Aspekty międzynarodowe (jeśli są)	
Język wykładowy	Polski
Cel przedmiotu Opis zakładanych kompetencji i umiejętności, jakie student nabywa w wyniku zaliczenia przedmiotu. Uwaga: maksymalna objętość tekstu to 3 linie standardowej strony A4	Sprawność w rozwiązywaniu praktycznych problemów z wykorzystaniem metod statystycznych i pakietu R. Znajomość szerokiej gamy modeli statystycznych oraz umiejętność doboru stosownej procedury statystycznej i jej implementacji dla rzeczywistych danych.
Treść przedmiotu	
<ol style="list-style-type: none"> Ogólne zapoznanie z pakietem R: metody wczytywania danych, rodzaje zmiennych, operacje na zmiennych, generowanie danych. Wstępna analiza danych. Wyznaczanie prawdopodobieństw i kwantyle dla wybranych rozkładów. Testy parametryczne dla jednej populacji: zastosowanie testów dla średniej, wariancji i wskaźnika struktury. Wyznaczanie niezbędnej ilości pomiarów potrzebnych do przeprowadzenia zadanego testu. Testy parametryczne dla dwóch populacji: zastosowanie testów do porównywania średnich, wariancji i wskaźników struktury. Wyznaczanie niezbędnej ilości pomiarów potrzebnych do przeprowadzenia danego testu. Analiza danych jakościowych: testowanie zgodności i sprawdzanie niezależności. Regresja liniowa jednokrotna: dopasowywanie modelu do danych, sprawdzanie czy dopasowany model dobrze opisuje dane, obserwacje wpływowe i odstające, przekształcanie zmiennych, metoda ważonych najmniejszych kwadratów, prognozowanie przy użyciu modelu regresji. Regresja liniowa wielokrotna: przekształcanie zmiennych, współliniowość zmiennych objaśniających, dobór zmiennych objaśniających do modelu (częściowy test F, kryteria służące do wyboru najlepszego modelu, metoda dołączania, eliminacji i selekcji krokowej). Jednoczynnikowa i dwuczynnikowa analiza wariancji: analiza wariancji jako szczególny przypadek regresji liniowej, wykresy średnich w grupach, sprawdzanie czy spełnione są założenia modelu, testy w modelu analizy wariancji, interakcje w dwuczynnikowej analizie wariancji. 	
Spis zalecanych lektur	
LP.	Autor, Tytuł, Wydawnictwo,
1.	P. Dalgaard, „Introductory Statistics with R”, Springer, 2008
2.	J. Koronacki, J. Mielniczuk, „Statystyka dla studentów kierunków technicznych i przyrodniczych”, Wydawnictwa Naukowo-Techniczne, Warszawa, 2006
3.	P. Biecek, „Przewodnik po pakiecie R”, Oficyna Wydawnicza GIS, Wrocław, 2008

CSZ sylabus 2014/2015/ 2

Wykład współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



4.	J.J. Faraway „Practical Regression and ANOVA Using R”, www.stat.lsa.umich.edu/~faraway/book
----	--

Metody oceny (ocena, egz. pisemny, egz. ustny, projekt)	Praca podczas zajęć – rozwiązywanie projektów.
--	--

Uwagi dodatkowe	Zajęcia odbędą się, jeżeli zapisze się co najmniej 20 osób.
------------------------	---

Tabela 1. Efekty kształcenia

Numer (symbol)	Efekty kształcenia słuchacz, który zaliczył przedmiot, potrafi	Sposób weryfikacji osiągnięcia efektu
WIEDZA		
AZETC_W1	Zna szeroką gamę testów statystycznych służących do analizy jednej populacji oraz do porównywania dwóch populacji.	Projekt
AZETC_W2	Zna metody graficzne i testy służące do analizy danych jakościowych.	Projekt
AZETC_W3	Zna i wie jak dopasować do danych następujące modele: model regresji liniowej jednokrotnej, regresji liniowej wielokrotnej oraz jednoczynnikowej i dwuczynnikowej analizy wariancji.	Projekt
UMIEJĘTNOŚCI		
AZETC_U1	Potrafi dobrać odpowiednią procedurę statystyczną do rozwiązania konkretnego, praktycznego problemu.	Projekt
AZETC_U2	Mając rzeczywiste dane, potrafi zastosować do nich wybraną procedurę statystyczną a następnie zweryfikować adekwatność modelu skonstruowanego w oparciu o wybraną procedurę.	Projekt
AZETC_U3	Przeprowadzenie formalnych testów podczas statystycznej analizy danych umie poprzedzić wstępną analizą wykorzystującą metody graficzne.	Projekt
AZETC_U4	Umie sprawnie posługiwać się pakietem R podczas przeprowadzania analizy danych.	Projekt
KOMPETENCJE		
AZETC_K1	Rozumie konieczność diagnozowania modelu, użytego do opisu rzeczywistych danych, i wprowadzania w nim stosownych modyfikacji.	Obserwacja na zajęciach
AZETC_K2	Rozumie, że automatyczne użycie procedur statystycznych, bez wniknięcia w istotę i charakter danych, grozi wyciągnięciem błędnych wniosków.	Obserwacja na zajęciach.

CSZ sylabus 2014/2015/ 3

Wykład współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPOJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



CSZ sylabus 2014/2015/ 4

Wykład współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

