



Nowoczesna Analiza Danych

Numer usługi 2025/07/21/135462/2893462

8 600,00 PLN brutto
8 600,00 PLN netto
41,35 PLN brutto/h
41,35 PLN netto/h

SZKOŁA GŁÓWNA
GOSPODARSTWA
WIEJSKIEGO W
WARSZAWIE

★★★★★ 4,8 / 5
19 ocen

📍 zdalna w czasie rzeczywistym
📖 Studia podyplomowe
🕒 208 h
📅 01.10.2025 do 30.09.2026

Informacje podstawowe

Kategoria	Inne / Edukacja
Grupa docelowa usługi	Studia przeznaczone są dla osób zainteresowanych szeroko pojętą analizą danych. Materiał wszystkich przedmiotów realizowany jest od podstaw, a więc nawet osoby bez przygotowania mogą uczestniczyć w zajęciach. Studenci poznają metody przygotowywania i selekcji analizowanych danych. Pozyskują umiejętność wyboru odpowiedniej metody analitycznej, skonfigurowania jej i zastosowania. W trakcie studiów studenci poznają różne środowiska do analizy danych, m.in. MS Excel, Pakiet R, Power BI, Python, SQL Server, Oracle Data Mining. Absolwenci studiów będą posiadać umiejętności pozwalające na precyzyjną analizę danych z punktu widzenia zasad statystyki, fachowe szacowanie błędów oraz określania istotności uzyskanych wyników. Nabędą wiedzę i umiejętności pozwalające na posługiwanie się współczesnymi narzędziami analizy danych, pozwalającymi na tworzenie dedykowanych modeli danych oraz ich efektywną analizę.
Minimalna liczba uczestników	15
Maksymalna liczba uczestników	40
Data zakończenia rekrutacji	30-09-2025
Forma prowadzenia usługi	zdalna w czasie rzeczywistym
Liczba godzin usługi	208

Cel

Cel edukacyjny

Absolwent będzie

- posiadał wiedzę matematyczną niezbędną do efektywnej analizy danych i wnioskowania;
- potrafił korzystać z wybranych narzędzi dedykowanych zadaniom Data Science;
- potrafił zaprojektować w wybranym narzędziu odpowiednie do zadania metody transformacji danych, ich selekcji i czyszczenia;
- potrafił samodzielnie przeprowadzić analizę danych, dobierając odpowiednią metodę, weryfikując procedurę oraz wyciągając wnioski.

Efekty uczenia się oraz kryteria weryfikacji ich osiągnięcia i Metody walidacji

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
Absolwent zna i rozumie podstawowe pojęcia z matematyki w zakresie niezbędnym do zrozumienia zaawansowanych zagadnień technik informatycznych.	egzamin końcowy	Test teoretyczny
Absolwent zna i rozumie zastosowania informatyki w naukach ekonomicznych i przyrodniczych	egzamin końcowy	Test teoretyczny
Absolwent potrafi zredagować, przeanalizować, a następnie zrealizować wymagania w przedsięwzięciach związanych z analizą, eksploracją i prezentacją danych o różnorodnym pochodzeniu	zadania wykonywane w ramach studiów	Analiza dowodów i deklaracji
	egzamin końcowy	Test teoretyczny
Absolwent jest gotów do odpowiedzialności za wspólnie realizowane zadania, związane z pracą zespołową.	zadania do samodzielnego lub grupowego wykonania	Analiza dowodów i deklaracji

Kwalifikacje

Kompetencje

Usługa prowadzi do nabycia kompetencji.

Warunki uznania kompetencji

Pytanie 1. Czy dokument potwierdzający uzyskanie kompetencji zawiera opis efektów uczenia się?

TAK

Pytanie 2. Czy dokument potwierdza, że walidacja została przeprowadzona w oparciu o zdefiniowane w efektach uczenia się kryteria ich weryfikacji?

TAK

Pytanie 3. Czy dokument potwierdza zastosowanie rozwiązań zapewniających rozdzielenie procesów kształcenia i szkolenia od walidacji?

Program

Opis przedmiotów

Studia rozpoczynają się od wykładu wprowadzającego w tematykę studiów. Następnie krok po kroku wprowadzane są podstawy statystycznej analizy danych oraz statystyki matematycznej. Poparte są one przykładami opartymi na rzeczywistych danych. Te i analogiczne dane są również analizowane w środowisku R, a w drugim semestrze w środowisku Python. Niezależnie, uczestnicy studiów uczą się pracy z relacyjnymi bazami danych, w szczególności tworzenia rozbudowanych raportów w oparciu o język SQL. Drugi semestr studiów poświęcony jest przede wszystkim dwóm podstawowym obecnie technologiom: chmurom obliczeniowym oraz zaawansowanym metodom analizy danych. Te ostatnie obejmują metody modelowania oraz analizy danych, jak również prezentacji i wymiany wyników.

Podstawy statystycznej analizy danych

W analizie danych niezbędna jest znajomość podstawowych zagadnień statystyki opisowej i matematycznej. Ilościowy opis analizowanych danych i ich prezentację graficzną umożliwia znajomość statystyki opisowej. Do wnioskowania statystycznego na podstawie badanych danych niezbędna jest podstawowa wiedza z zakresu analizy błędów, która z kolei stanowi podstawę statystyki matematycznej.

1. Ilościowy opis danych statystycznych, miary średnie, wariancja, odchylenie standardowe, kwantyle.
2. Graficzna prezentacja danych, wykres pudełkowy, histogram.
3. Skośność i kurtoza rozkładu empirycznego.
4. Błąd statystyczny, błąd wartości średniej.
5. Rozkład Gaussa i rozkłady z próby.
6. Estymatory, testy statystyczne, przedziały ufności.
7. Analiza dwóch zmiennych, regresja liniowa. Regresja liniowa wielu zmiennych.

W ramach ćwiczeń, przykłady analizy danych prowadzone są w programie Excel oraz w Pakiecie R.

Wstęp do analizy wielowymiarowej

1. Istota danych wielowymiarowych i podstawowe problemy związane z ich analizą.
2. Transformacja danych oraz redukcja wymiaru.
3. Podstawowe metody analizy danych wielowymiarowych, techniki Data Mining i Machine Learning.

MS SQL Server

Microsoft SQL Server to jeden z najpopularniejszych serwerów bazodanowych na świecie. Na jego przykładzie słuchacze studiów poznają podstawy relacyjnych baz danych oraz język SQL.

1. Wstęp do relacyjnych baz danych i języka SQL.
2. Wydobywanie danych z bazy danych i formatowanie wyników.
3. Filtrowanie rekordów.
4. Grupowanie danych.
5. Wydobywanie danych z wielu tabel.

Pakiet R

R jest językiem programowania dystrybuowanym na licencji open-source. Służy do analiz statystycznych i wizualizacji danych. Zdobył dużą popularność w środowiskach naukowych i akademickich, ale jest też wykorzystywany w wielu firmach i instytucjach: Facebook, Google, Merck, Altera, Pfizer, LinkedIn, Shell, Novartis, Ford, Mozilla, czy Twitter. Jest wciąż rozbudowywany, pojawiają się nowe specjalistyczne pakiety.

W ramach przedmiotu **Podstawy Statystycznej Analizy Danych** (6 godz.).

1. Wstęp do środowiska R, instalacja pakietów, wczytywanie i zapisywanie danych, struktury danych w R, przekształcanie danych, filtrowanie.
2. Elementy programowania: tworzenie skryptów obliczeniowych, obliczenia z wykorzystaniem pętli, budowa własnych funkcji.
3. Prezentowanie danych i wyników: wykresy słupkowe, kołowe, histogramy, wykresy pudełkowe (boxplot).
4. Metody statystyczne: statystyczny opis danych, przedziały ufności, testy parametryczne, testy zgodności i niezależności.

W ramach przedmiotu **Pakiet R** (18 godz.).

1. Regresja liniowa jednej i wielu zmiennych, kryteria informacyjne, punkty odstające i wpływowe, metody graficzne.
2. Regresja logistyczna, drzewa klasyfikacyjne, lasy losowe.
3. Metody analizy skupień, w tym metoda k-średnich.
4. Metoda głównych składowych, metoda wektorów nośnych.
5. Drzewa regresyjne.

Power BI

Power BI to aplikacja firmy Microsoft typu desktop przeznaczona do pracy w systemie Windows. Zdobywa coraz większą popularność z powodu swojej kompleksowości (umożliwia tzw. analizę typu self-service) oraz współpracy z chmurą. Aplikacja umożliwia łączenie się całym spektrum różnych źródeł danych, pobieranie ich i modelowanie. Posiada wbudowanych coraz więcej elementów hurtowni danych, również tych, związanych z technikami Data Mining. Pozwala na projektowanie i udostępnianie raportów i wyników analiz.

1. Wstęp do architektury Power BI.
2. Pozyskiwanie i transformowanie danych.
3. Tworzenie modeli danych.
4. Analiza danych.
5. Tworzenie elementów wizualnych i raportów.

Python

Przedmiot ten obejmuje wszystkie najważniejsze zagadnienia związane z językiem programowania obiektowo-skryptowym Python. Python to obecnie jeden z najpopularniejszych języków programowania, który swoją elastycznością stał się jednym z najważniejszych rozwiązań również do procesów ETL oraz ELT. Wraz z rozwiązaniami chmurowymi stawowi on dopełnienie w rozwiązaniach business intelligence.

1. Wstęp do programowania w języku Python. Środowisko, instalacja oraz zarządzanie pakietami. Narzędzia do programowania (IDE), metody wersjonowania i testowania kodu.
2. Omówienie typów danych, funkcji, operatorów, itd.
3. Pętle i instrukcje warunkowe.
4. Konstrukcja metod, funkcji i klas.
5. Omówienie wykorzystania najważniejszych pakietów dostępnych w środowisku Python wraz z ich przykładowym zastosowaniem.
6. Wykorzystanie Pythona do wydobywania i analizy danych.

MS SQL Business Intelligence

Microsoft SQL Server Business Intelligence to kompletna platforma dla środowisk inteligencji biznesowej. Składa się ona z serwera baz danych, środowiska Integration Services do tworzenia rozwiązań ETL (Extract – Transform – Load) służących do dystrybucji i transformowania danych, Analysis Services do projektowania i tworzenia wydajnych hurtowni danych opartych na architekturze wielowymiarowej lub tabelarycznej oraz Reporting Services do projektowania, tworzenia i dystrybucji raportów.

1. Wstęp do środowiska MS SQL Server BI.
2. Tworzenie pakietów do transferowania i transformowania danych.
3. Tworzenie elementów hurtowni danych: wymiary, kostki, obliczenia, itp.
4. Tworzenie hurtowni danych.
5. Tworzenie modeli Data Mining i analiza danych.

6. Projektowanie i tworzenie raportów.

Oracle Data Mining

Firma Oracle jest największym dostawcą rozwiązań bazodanowych na świecie. Oferuje nie tylko serwery bazodanowe, ale również cały szereg aplikacji, w tym do analizy danych. W tej części studiów słuchacze zapoznają się z bazami danych Oracle i podstawami języka PL/SQL. Główną częścią przedmiotu będzie przygotowanie danych i ich analiza przy pomocy technik Data Mining oraz Machine Learning (uczenie maszynowe).

1. Wstęp do baz danych Oracle i PL/SQL.
2. Integracja Oracle z pakietami Python i R.
3. Techniki Data Mining i Machine Learning.
4. Analiza wspólnego koszyka.
5. Wykrywanie anomalii.
6. Wdrażanie rozwiązań Data Mining i Machine Learning.

Rozwiązania chmurowe

Rozwiązania chmurowe omawiane są na przykładzie technologii Apache Hadoop. Jest to zestaw bibliotek i narzędzi służących do przechowywania, przetwarzania oraz analizowania dużych zbiorów danych. Jest on ściśle powiązany z takimi pojęciami jak Cloud Computing, Big Data, Machine Learning, czy Artificial Intelligence. Apache Hadoop pozwala przechowywać i przetwarzać duże ilości danych przy pomocy wielu zadań wsadowych. Do przetwarzania danych na Hadoop można jeszcze użyć Apache Spark, czyli platformy opartej na klastrach. Spark dostarczą nowe możliwości programiście, których sam Hadoop nie posiada. Oba te rozwiązania mogą używać programu Hive, który tworzy interfejs dostępu do danych HiveQL, który bazuje na języku SQL. HBase jest przykładem aplikacji uruchomionej na systemie plików HDFS, który jest częścią Hadoopa. Baza HBase jest modelowana w oparciu o Google Bigtable czyli nie-relacyjnej bazie danych.

1. Wprowadzenie do tematyki Big Data.
2. Komponenty środowiska Hadoop. HDFS – rozproszone przechowywanie danych.
3. Nierelacyjne bazy danych. Hive – omówienie działania MapReduce i ćwiczenia w operowaniu językiem HQL.
4. Apache Spark – komponenty i charakterystyka przetwarzania danych.
5. Wprowadzenie i ćwiczenia w użyciu biblioteki PySpark.
6. Tworzenie, wdrażanie i optymalizacja aplikacji sparkowych.

Harmonogram

Liczba przedmiotów/zajęć: 31

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
1 z 31 Wykład wprowadzający	-	25-10-2025	09:00	10:30	01:30
2 z 31 Podstawy statystycznej analizy danych	-	25-10-2025	10:45	17:45	07:00
3 z 31 Podstawy statystycznej analizy danych	-	26-10-2025	09:00	16:00	07:00

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
4 z 31 Podstawy statystycznej analizy danych	-	15-11-2025	09:00	17:45	08:45
5 z 31 Podstawy statystycznej analizy danych	-	16-11-2025	09:00	16:00	07:00
6 z 31 Pakiet R	-	29-11-2025	09:00	14:15	05:15
7 z 31 Wstęp do analizy wielowymiarowej	-	29-11-2025	14:30	17:45	03:15
8 z 31 Wstęp do analizy wielowymiarowej	-	30-11-2025	09:00	10:30	01:30
9 z 31 SQL Server	-	30-11-2025	10:45	17:45	07:00
10 z 31 SQL Server	-	13-12-2025	09:00	17:45	08:45
11 z 31 SQL Server	-	14-12-2025	09:00	14:15	05:15
12 z 31 Pakiet R	-	14-12-2025	14:30	17:45	03:15
13 z 31 Pakiet R	-	10-01-2026	09:00	17:45	08:45
14 z 31 Pakiet R	-	11-01-2026	09:00	12:15	03:15
15 z 31 Power BI	-	11-01-2026	12:45	17:45	05:00
16 z 31 Power BI	-	24-01-2026	09:00	16:00	07:00
17 z 31 Power BI	-	25-01-2026	09:00	16:00	07:00
18 z 31 Python	-	14-02-2026	09:00	17:45	08:45
19 z 31 Python	-	14-03-2026	09:00	14:15	05:15
20 z 31 Business Intelligence	-	14-03-2026	14:30	17:45	03:15
21 z 31 Business Intelligence	-	15-03-2026	09:00	16:00	07:00

Przedmiot / temat zajęć	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
22 z 31 Business Intelligence	-	28-03-2026	09:00	17:45	08:45
23 z 31 Business Intelligence	-	29-03-2026	09:00	12:15	03:15
24 z 31 Oracle Data Mining	-	29-03-2026	12:45	16:00	03:15
25 z 31 Oracle Data Mining	-	18-04-2026	09:00	16:00	07:00
26 z 31 Oracle Data Mining	-	19-04-2026	09:00	16:00	07:00
27 z 31 Oracle Data Mining	-	09-05-2026	09:00	12:15	03:15
28 z 31 Rozwiązania Chmurowe	-	09-05-2026	12:45	16:00	03:15
29 z 31 Rozwiązania Chmurowe	-	10-05-2026	09:00	16:00	07:00
30 z 31 Rozwiązania Chmurowe	-	13-06-2026	09:00	16:00	07:00
31 z 31 Rozwiązania Chmurowe	-	14-06-2026	09:00	12:15	03:15

Cennik

Cennik

Rodzaj ceny	Cena
Koszt przypadający na 1 uczestnika brutto	8 600,00 PLN
Koszt przypadający na 1 uczestnika netto	8 600,00 PLN
Koszt osobogodziny brutto	41,35 PLN

Prowadzący

Liczba prowadzących: 0

Brak wyników.

Informacje dodatkowe

Informacje o materiałach dla uczestników usługi

Treści są prezentowane w trakcie wykładów w formie plików pdf. W trakcie ćwiczeń uczestnicy otrzymują materiały dydaktyczne w różnych formatach.

Warunki uczestnictwa

Słuchaczami studiów mogą być osoby, które ukończyły studia wyższe i posiadają dyplom: licencjata, inżyniera, magistra lub magistra inżyniera. Ponad to są to osoby, które wniosły wszystkie, przewidziane regulaminem studiów opłaty (wpisowe + opłata za jeden lub oba semestry studiów).

Warunki techniczne

Zajęcia odbywają się za pośrednictwem platformy MS Teams. Wymagane jest posiadanie komputera z dostępem do internetu. Niezbędne oprogramowanie uczestnicy będą instalować przed zajęciami zgodnie ze wskazówkami wykładowców.

Kontakt



Tomasz Minkowski

E-mail spnad@sggw.edu.pl

Telefon (+48) 695 797 913