



Karta przedmiotu

Nazwa i kod przedmiotu	Statystyka i analiza danych, PG_00060847						
Kierunek studiów	Technologia chemiczna						
Data rozpoczęcia studiów	październik 2023 r.	Rok akademicki realizacji przedmiotu			2023/2024		
Poziom kształcenia	I stopnia - inżynierskie	Grupa zajęć			Grupa zajęć obowiązkowych z zakresu kierunku studiów		
Forma studiów	stacjonarne	Sposób realizacji			na uczelni		
Rok studiów	1	Język wykładowy			polski		
Semestr studiów	2	Liczba punktów ECTS			2.0		
Profil kształcenia	ogólnoakademicki	Forma zaliczenia			zaliczenie		
Jednostka prowadząca	Wydział Chemiczny -> Katedra Chemii Fizycznej						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot	dr hab. inż. Adam Kloskowski					
	Prowadzący zajęcia z przedmiotu	dr hab. inż. Adam Kloskowski					
Formy zajęć i metody nauczania	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	15.0	0.0	0.0	30
	W tym liczba godzin zajęć na odległość: 0.0						
Aktywność studenta i liczba godzin pracy	Aktywność studenta	Udział w zajęciach dydaktycznych, objętych planem studiów		Udział w konsultacjach		Praca własna studenta	RAZEM
	Liczba godzin pracy studenta	30		5.0		15.0	50
Cel przedmiotu	Student będzie umiał wyjaśnić podstawowe pojęcia z zakresu statystyki. Student po odbyciu laboratoriów potrafi wykorzystać oprogramowanie Excel oraz Python (Orange Data Mining). do analizy danych eksperymentalnych Po ukończeniu kursu student powinien:  1) w sposób biegły posługiwać się zaawansowanymi funkcjami programu Excel  2) znać podstawy posługiwania się językiem Python  3) potrafić tworzyć modele matematyczne (regresyjne) służące rozwiązywaniu podstawowych problemów technologicznych						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	<p>[K6_W01] ma wiedzę w zakresie matematyki, obejmującą rozwiązywanie równań i nierówności zawierających funkcje elementarne, rachunek różniczkowy i całkowy, elementy analizy wektorowej, statystyki, optymalizacji i metod numerycznych, ma podstawową wiedzę w zakresie wybranych działów fizyki, przydatną do opisu i analizy procesów technologicznych</p>	<p>Student posiada wiedzę z zakresu matematyki niezbędną dla właściwego opisu statystycznego zestawów danych. Student posiada wiedzę z zakresu fizyki niezbędną w procesie analizy danych dotyczących zagadnień technicznych oraz technologicznych.</p>	<p>[SW1] Ocena wiedzy faktograficznej [SW2] Ocena wiedzy zawartej w prezentacji [SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym</p>
	<p>[K6_U01] potrafi pozyskiwać informacje z literatury, baz danych oraz innych, właściwie dobranych źródeł, także w języku angielskim; potrafi integrować uzyskane informacje, dokonywać ich interpretacji, a także wyciągać wnioski, formułować i uzasadniać opinie</p>	<p>Student potrafi przygotować interesujący sposób prezentację danych statystycznych. Biegłe posługuje się narzędziami służącymi do tworzenia wykresów i prezentacji zależności między zmiennymi. Student posiada umiejętność analizy informacji w kontekście wpływu podejmowanych decyzji na środowisko. Ma świadomość odpowiedzialności za podejmowane decyzje. Potrafi pracować w grupie jak i indywidualnie oraz jest świadomy konieczności dotrzymania założonych terminów</p>	<p>[SU1] Ocena realizacji zadania [SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu [SU5] Ocena umiejętności zaprezentowania wyników realizacji zadania</p>
Treści przedmiotu	<p>W trakcie zajęć omawiane będą zagadnienia wprowadzające studenta w podstawowe pojęcia z zakresu statystyki mające na celu wyposażać studenta w narzędzia niezbędne na późniejszych etapach kształcenia z zakresu opracowania wyników pomiarów, planowania i przeprowadzania eksperymentów oraz metrologii. W ramach wykładów studenci zapoznają się z podstawami opisu statystycznego w zakresie pojęć oraz narzędzi obliczeniowych (z wprowadzeniem do metrologii). W ramach laboratorium rozwiązują zadania praktyczne związane z wykorzystaniem narzędzi statystycznych oraz odpowiednio dobranego oprogramowania (Excel, Python) w kontekście ogólnym jak i technicznym. Treści omawiane w ramach wykładu oraz laboratorium obejmują następujące obszary:</p> <p>Opis statystyczny zbioru danych</p> <ul style="list-style-type: none"> <li>- statystyki jednowymiarowej zmiennej losowej</li> <li>- błąd bezwzględny, błąd względny, precyzja, dokładność</li> <li>- określanie niepewności przyrządów pomiarowych, niepewność etapu kalibracji metod eksperymentalnych, metoda propagacji niepewności, metody szacowania niepewności (typu A i B), diagram Ishikawy</li> <li>- poprawny zapis wyników pomiarów eksperymentalnych wraz z niepewnością i jednostką, wprowadzenie do pojęcia spójności pomiarowej,</li> <li>- miary położenia (ze szczególnym uwzględnieniem takich jak średnia arytmetyczna, geometryczna, dominanta, mediana, kwantyle); miary rozproszenia (ze szczególnym uwzględnieniem takich jak odchylenie standardowe, współczynnik zmienności, rozstęp); miary asymetrii,</li> <li>- metody reprezentacji danych: histogram, diagram pudełkowy</li> <li>- rozkład normalny, t-Studenta, chi-kwadrat, Poissona (małych liczb), Boltzmanna, jednostajny, trójkątny, skośność rozkładu, dystrybuanta rozkładu, centralne twierdzenie graniczne</li> <li>- wyznaczenie przedziału ufności</li> <li>- pojęcia powszechnie stosowane w praktyce laboratoryjnej: powtarzalność, odtwarzalność, dokładność, poprawność, liniowość, zakres pomiarowy, czułość, kalibracja, szum i granica wykrywalności</li> </ul> <p>Wnioskowanie statystyczne weryfikacja hipotez</p> <ul style="list-style-type: none"> <li>- pojęcia: hipoteza zerowa, hipoteza alternatywna, poziom istotności, obszar krytyczny testu, testy parametryczne i nieparametryczne</li> <li>- procedury wnioskowania statystycznego</li> <li>- rodzaje błędów: błędy systematyczne, błędy przypadkowe, błędy grube,</li> <li>- szacowanie prawdopodobieństwa zajścia zdarzenia</li> <li>- testy statystyczne: Q-Dixona, Grubbsa, F-Snedecora, t-Studenta, Aspina-Welcha, Cochran-Coxa,</li> <li>- reguła 3 sigm</li> <li>- wartość p</li> </ul> <p>Analiza danych</p> <ul style="list-style-type: none"> <li>- przedstawianie danych dwuwymiarowych: wykres punktowy, liniowy</li> <li>- przedstawianie przedziału ufności oraz wartości niepewności na wykresie</li> <li>- korelacja, współczynnik korelacji Pearsona i Spearmana</li> <li>- regresja liniowa, zależności nieliniowe</li> <li>- wstęp do analizy danych wielowymiarowych, regresja wielokrotna</li> <li>- wstęp do analizy big data: przetwarzanie i czyszczenie zbiorów danych, określanie i przewidywanie wzorców i relacji w zbiorach danych</li> </ul> <p>Zastosowania metod i narzędzi statystycznych</p> <ul style="list-style-type: none"> <li>- Zastosowania w analizie danych eksperymentalnych (kalibracja procedury, metody typu QSAR, analiza skupień)</li> <li>- Walidacja metody pomiarowej</li> <li>- Elementy optymalizacji eksperymentalnej</li> <li>- Planowanie doświadczeń z uwzględnieniem planów czynnikowych i minimalnych</li> <li>- Statystyczne kryteria oceny miarodajności wyników i porównywania metod eksperymentalnych</li> </ul>		
Wymagania wstępne i dodatkowe	Znajomość matematyki w zakresie podstawowym		

Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Laboratorium – zadanie problemowe	50.0%	50.0%
	Wykład -test	50.0%	50.0%
Zalecana lista lektur	Podstawowa lista lektur	Wojciech Hyk, Zbigniew Stojek, Analiza statystyczna w laboratorium badawczym, PWN, Warszawa 2019 - Andrzej Balicki, Wiesława Makać, Metody wnioskowania statystycznego, Wydawnictwo UG, Gdańsko 2006 - Felix Zumstein, Python i Excel. Nowoczesne środowisko do automatyzacji i analizy danych, Helion, Warszawa 2021] - James Miller, Jane Miller, Statystyka i chemometria w chemii analitycznej, PWN, Warszawa 2016 - YouTube: Orange Data Mining tutorials [dostępne online]	
	Uzupełniająca lista lektur	- P. Konieczka Ocena i kontrola jakości wyników pomiarów analitycznych, WNT, Warszawa 2007 - J. Mazerski Podstawy chemometrii, Wydawnictwo PG, Gdańsk 2004 - A. Navlani, A. Fandango, I. Idris, Python i praca z danymi. Przetwarzanie, analiza, modelowanie i wizualizacja, Helion, Warszawa 2022 - Joel Grus, Data science	
	Adresy eZasobów	Adresy na platformie eNauczanie:	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	Przykładowe zagadnienia teoretyczne: - Ile cyfr przedstawić w mierzonym wyniku? Jakimi metodami szacuje się niepewność pomiaru? - Co to jest współczynnik R <sup>2</sup> ? Czy może on przyjmować wartości ujemne i jeśli tak to kiedy się tak dzieje? - Co to jest korelacja? Czy wysokie skorelowanie zmiennych w modelu regresji jest korzystne? - Niepewność a błąd - omów znaczenie tych sformułowań. - Wyjaśnij pojęcia: precyzja, dokładność, czułość, specyficzność. - Omów dostępne w programie Excel narzędzia do wyznaczenia odchylenia standardowego. Przedyskutuj różnice między nimi (np. podając wzory według której każda z funkcji je oblicza) - Omów wybrane metody identyfikacji obserwacji oddalonych. - Na czym polega linearyzacja? Przykładowe zagadnienia obliczeniowe: - Używając metody regresji wielokrotnej (model liniowy) znajdź zależność między toksycznością a podanymi w tabeli wartościami zmiennych dla podanych zestawów przykładów składników kosmetyków. Z użyciem regresji liniowej wyznacz wartość pH próbki wody z rzeki, korzystając z wyników kalibracji pH-metru elektrochemicznego zamieszczonych w tabeli - Oceń czy badane warunki fermentacji mają wpływ na skład (wydajność otrzymywania itp.) wina z użyciem wybranego testu statystycznego. - Na podstawie zbioru danych dokonać oceny dokładności i precyzji techniki pomiarowej		
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy		