

KARTA KURSU (realizowanego w specjalności)

Data Science

(nazwa specjalności)

Nazwa	Analiza danych z językiem SQL
Nazwa w j. ang.	<i>Data Analysis with SQL</i>

Koordynator	dr Roman Czapla	Zespół dydaktyczny
Punktacja ECTS*	st. stacjonarne: 4 st. niestacjonarne: 4	

Opis kursu (cele kształcenia)

Celem kursu jest nabycie praktycznych umiejętności analizy danych z wykorzystaniem języka SQL w kontekście projektów Data Science. Studenci uczą się wykorzystywać SQL do eksploracji, transformacji i przygotowania danych do dalszych analiz statystycznych i modelowania w Pythonie. Kurs obejmuje zarówno klasyczne zapytania analityczne, jak i zaawansowane techniki, takie jak funkcje okienkowe, analiza kohortowa, przetwarzanie danych tekstowych oraz wykrywanie anomalii.

Szczególny nacisk położony jest na integrację SQL z językiem Python, co pozwala studentom na budowanie kompletnych przepływów analitycznych, łączących zapytania do baz danych z obliczeniami i wizualizacjami w środowisku programistycznym.

Warunki wstępne

Wiedza	Student zna podstawy analizy danych oraz orientuje się w strukturze relacyjnych baz danych i podstawowych operacjach zapytań SQL.
Umiejętności	Student potrafi korzystać z języka Python do pracy z danymi (np. za pomocą bibliotek pandas, matplotlib) oraz uruchamiać proste zapytania SQL w środowisku programistycznym.
Kursy	

Efekty uczenia się

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Wiedza	Po zakończeniu kursu: W01: Student zna możliwości języka SQL w kontekście analizy danych i eksploracji zbiorów informacji. W02: Student zna zaawansowane konstrukcje zapytań SQL, w tym funkcje okienkowe, agregacje i operacje na danych czasowych. W03: Student rozumie różnice i zależności między analizą danych w SQL i Pythonie. W04: Student zna podstawowe techniki przygotowania cech (feature engineering) i analizy kohortowej. W05: Student zna metody wykrywania anomalii i pracy z danymi tekstowymi w SQL.	

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Umiejętności	Po zakończeniu kursu: U01: Student potrafi formułować złożone zapytania SQL do celów analitycznych, w tym z użyciem CTE, funkcji warunkowych i okienkowych. U02: Student potrafi analizować dane czasowe, tworzyć kohorty i wskaźniki retencji z wykorzystaniem SQL. U03: Student potrafi przekształcać i agregować dane w SQL oraz wizualizować wyniki analiz w Pythonie. U04: Student potrafi przygotować dane tekstowe i liczbowe do dalszego przetwarzania w projektach analitycznych. U05: Student potrafi zintegrować SQL z Pythonem w środowisku Jupyter Notebook do realizacji złożonych zadań analitycznych.	

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Kompetencje społeczne	Po zakończeniu kursu: K01: Student potrafi pracować w zespole nad realizacją zadań analitycznych w oparciu o dane z relacyjnych baz danych. K02: Student rozumie znaczenie poprawności i przejrzystości zapytań oraz odpowiedzialnego przetwarzania danych.	

Studia stacjonarne

Organizacja							
Forma zajęć	Wykład (W)	Ćwiczenia w grupach					
		A	K	L	S	P	E
Liczba godzin	10			40			

Studia niestacjonarne

Organizacja							
Forma zajęć	Wykład (W)	Ćwiczenia w grupach					
		A	K	L	S	P	E
Liczba godzin	5			25			

Opis metod prowadzenia zajęć

Zajęcia mają charakter praktyczny i są prowadzone w formie ćwiczeń w środowisku Jupyter Notebook. Studenci pracują z bazami danych SQL i uczą się budować zapytania analityczne oraz integrować je z językiem Python. Każde zajęcia obejmują zestaw zadań praktycznych ukierunkowanych na konkretne zagadnienia analizy danych – od agregacji po analizę kohortową i przetwarzanie danych tekstowych. Zajęcia kończą się realizacją projektu, który łączy różne techniki pracy z danymi w SQL i Pythonie.

Formy sprawdzania efektów uczenia się

	E – learning	Gry dydaktyczne	Ćwiczenia w szkole	Zajęcia terenowe	Praca laboratoryjna	Projekt indywidualny	Projekt grupowy	Udział w dyskusji	Referat	Praca pisemna (esej)	Egzamin ustny	Egzamin pisemny	Zadania problemowe
W01													
W02													
U01													
U02													
K01													

Kryteria oceny	Osiągnięcie efektów kształcenia podanych powyżej uprawnia studentów do uzyskania oceny nie wyższej niż dostateczna. Warunkiem zaliczenia przedmiotu jest: - uzyskanie pozytywnej oceny z kolokwium sprawdzającego znajomość języka SQL w kontekście analizy danych, - zrealizowanie i oddanie projektu zaliczeniowego polegającego na przygotowaniu kompletnej analizy danych z wykorzystaniem SQL oraz Python, - aktywna praca na zajęciach oraz terminowe wykonywanie zadań praktycznych. Zaliczenie na ocenę dobrą lub bardzo dobrą otrzymuje student, który spełnia warunki oceny dostatecznej, a ponadto: - tworzy zoptymalizowane i przejrzyste zapytania SQL dostosowane do problemów analitycznych, - samodzielnie łączy różnorodne techniki SQL (funkcje okienkowe, analiza czasowa, kohortowa, tekstowa), - potrafi efektywnie zintegrować SQL z narzędziami Python w analizie danych i prezentacji wyników, - wykazuje inicjatywę i dbałość o dokumentację oraz jakość przygotowanego projektu.
----------------	--

Uwagi	
-------	--

Treści merytoryczne (wykaz tematów)

1. SQL w kontekście analizy danych i Data Science: zastosowanie SQL jako narzędzia eksploracji danych, integracja z językiem Python, porównanie przetwarzania danych w SQL i Pandas. 2. Zaawansowane zapytania i raportowanie: konstrukcje zapytań analitycznych, agregacje, przekształcenia danych, tworzenie raportów oraz eksport wyników do dalszej analizy w Pythonie. 3. Funkcje okienkowe: wykorzystanie funkcji analitycznych do analiz porównawczych, tworzenia rankingów i obserwacji zmian w czasie. 4. Analiza danych czasowych: operacje na datach i czasie, grupowanie po interwałach, tworzenie wskaźników czasowych i wizualizacja ich w Pythonie. 5. Analiza kohortowa i retencyjna: segmentacja użytkowników, obliczanie wskaźników retencji i interpretacja danych kohortowych z użyciem SQL i Python. 6. Tworzenie cech (feature engineering): przygotowanie danych wejściowych do modeli uczenia maszynowego, generowanie zmiennych opisujących aktywność i zachowania. 7. Analiza danych tekstowych: przetwarzanie i transformacja danych tekstowych w SQL, przygotowanie danych do dalszej analizy jakościowej lub ilościowej. 8. Wykrywanie anomalii i rozkładów: identyfikacja wartości odstających, analiza rozkładów statystycznych, integracja SQL z Pythonem w celu wizualizacji danych i statystyk.
--

Wykaz literatury podstawowej

1.

Wykaz literatury uzupełniającej

1.

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) – **studia stacjonarne**

Liczba godzin w kontakcie z prowadzącymi	Wykład	10
	Konwersatorium (ćwiczenia, laboratorium itd.)	40
	Pozostałe godziny kontaktu studenta z prowadzącym	
Liczba godzin pracy studenta bez kontaktu z prowadzącymi	Lektura w ramach przygotowania do zajęć	
	Realizacja zadań domowych (problemowych) po zapoznaniu się z niezbędną literaturą przedmiotu	
	Przygotowanie projektu lub prezentacji na podany temat (praca indywidualna lub w grupie)	
	Przygotowanie do egzaminu/zaliczenia	
Ogółem bilans czasu pracy		
Liczba punktów ECTS w zależności od przyjętego przelicznika		4

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) – **studia niestacjonarne**

Liczba godzin w kontakcie z prowadzącymi	Wykład	5
	Konwersatorium (ćwiczenia, laboratorium itd.)	25
	Pozostałe godziny kontaktu studenta z prowadzącym	
Liczba godzin pracy studenta bez kontaktu z prowadzącymi	Lektura w ramach przygotowania do zajęć	
	Realizacja zadań domowych (problemowych) po zapoznaniu się z niezbędną literaturą przedmiotu	
	Przygotowanie projektu lub prezentacji na podany temat (praca indywidualna lub w grupie)	
	Przygotowanie do egzaminu/zaliczenia	
Ogółem bilans czasu pracy		
Liczba punktów ECTS w zależności od przyjętego przelicznika		4