

## Lista zagadnień na egzamin inżynierski dla kierunku ANALIZA DANYCH

1. Rachunek zbiorów, rachunek zdań.
2. Indukcja matematyczna.
3. Funkcje elementarne i ich własności.
4. Pojęcie pochodnej funkcji w punkcie oraz jej interpretacja geometryczna i fizyczna
5. Podstawowe własności całki Riemanna, interpretacja geometryczna całki Riemanna.
6. Macierze. Podstawowe operacje na macierzach. Rząd i wyznacznik macierzy.
7. Rozwiązywanie układów równań liniowych.
8. Podstawowe elementy kombinatoryki.
9. Podstawowe pojęcia rachunku prawdopodobieństwa (prawdopodobieństwo zdarzenia, klasyczna definicja prawdopodobieństwa, prawdopodobieństwo geometryczne, zmienne losowe).
10. Prawdopodobieństwo warunkowe i niezależność zdarzeń (wzór na prawdopodobieństwo całkowite, wzór Bayesa, Schemat Bernoulliego). Niezależność zmiennych losowych.
11. Podstawowe rozkłady prawdopodobieństwa i ich własności (rozkład dwumianowy, geometryczny, Poissona wykładniczy, normalny).
12. Charakterystyki liczbowe zmiennych losowych oraz prób losowych (średnia, wartość oczekiwana, wariancja, odchylenie standardowe, współczynnik zmienności, skośność i kurtoza; kwantyle; współczynnik korelacji).
13. Podstawowe pojęcia statystyczne: próba losowa, statystyka, eksperyment statystyczny i jego cele. Własności estymatorów i metody ich konstrukcji. Testowanie hipotez statystycznych (błędy I i II rodzaju, poziom istotności i moc testu).
14. Sposoby prezentacji danych w arkuszach kalkulacyjnych. Zasady adresowania i odwoływania się do komórek w formułach arkuszy.
15. Makra w arkuszu kalkulacyjnym Excel (definicja, przeznaczenie, przykłady). Zdarzenia i elementy sterujące arkusza.
16. Średnie ruchome ważone wykładniczo. Sposoby badania trendów i generowania sygnałów transakcyjnych.
17. Wskaźniki techniczne mierzące siłę trendu ceny aktywów i generujące sygnały kupna/sprzedaży.
18. Symptomy techniczne zbliżającego się końca trendu wzrostowego/spadkowego.
19. Rodzaje danych statystycznych, metody ich opisu i prezentacji.
20. Analiza skupień, narzędzia matematyczne stosowane w tej analizie.
21. Reguły asocjacyjne, przykłady.
22. Struktura logiczna i funkcjonalna klasycznego komputera.
23. Arytmetyka stałopozycyjna i zmiennopozycyjna. Reprezentacja liczb w komputerze.
24. Iteracja i rekurencja – przykłady zastosowań oraz implementacji.
25. Cechy programowania strukturalnego i obiektowego.
26. Publikowanie w sieci Internet (HTML - własności, komponenty, przykłady, CSS - zasady tworzenia i stosowania stylów kaskadowych, zalety i ograniczenia).
27. Skrypty powłoki w systemie operacyjnym Linux. Automatyzacja podstawowych programów wiersza poleceń za pomocą skryptów powłoki.
28. Zaawansowana analityka biznesowa (klasyfikacja, grupowanie, odkrywanie reguł asocjacyjnych).

29. Przetwarzanie analityczne w trybie online (OLAP).
30. Podstawowe elementy teorii baz danych (definicje i własności bazy danych, model danych, schemat danych).
31. Podstawowe elementy relacyjnych baz danych (definicja i własności relacji, definicja i własności klucza, rodzaje kluczy, normalizacja).
32. Rodzaje kwerend w relacyjnych bazach danych.
33. Język SQL – własności, podział instrukcji i ich budowa, przykłady.
34. Pojęcie nierelacyjnej bazy danych (tzw. NoSQL). Funkcje, możliwości i różnice w porównaniu z bazami relacyjnymi. Przykłady takich baz. Przykłady wykorzystania.
35. Pojęcie dużego zbioru danych. Metody i narzędzia przetwarzania dużych zbiorów danych.
36. Metody uczenia nienadzorowanego i nadzorowanego – ich charakterystyka i przykłady zastosowań.
37. Algorytm budowy drzewa decyzyjnego.
38. Proces eksploracji danych (CRISP-DM).
39. Zasady i metody konstrukcji eksperymentów (randomizacja, schemat blokowy, schemat porównań parami)
40. Problem fałszywie pozytywnych wyników w wielokrotnym testowaniu i metody kontroli FDR
41. Statystyczne algorytmy uczące w analizie danych naukowych (PCA, lasso, bootstrap, testy permutacyjne).