Wymagania dotyczące projektu zaliczeniowego

Statystyka, Bazy Danych

Stworzenie repozytorium git - pierwszy tydzień, przesłanie linku

- wybór tematu do 27 października 2023

- I deadline: do końca zajęć (26 stycznia 2024) - koniecznie zaliczenie ćwiczeń (cała część z baz plus ćwiczenia z zajęć - statystyka)

- II deadline: koniec sesji (11 lutego 2024) - część statystyczna projektu

1. Przygotowanie odpowiedniej struktury bazy danych na wybrany temat. Przykłady tematów:
   1. Firma produkcyjna
   2. Firma sprzedażowa
   3. Sieć kin
   4. Obsługa ligi piłki
   5. Firma bukmacherska
   6. Firma sprzedająca nieruchomości
   7. Obsługa giełdy papierów wartościowych
   8. Prognoza pogody
   9. ...

Ważne, żeby były dane nadające się do obróbki statystycznej

1. Skrypt ładujący dane do bazy
   1. Może pobierać dane ze strony – np. odpalany co jakiś czas/ładowanie z plików csv, json. (deamon, cron table, itd)
   2. Sprawdzanie poprawności danych.
   3. Gromadzenie i archiwizacja przetworzonych/załadowanych danych do bazy.
2. Skrypt opisujący strukturę danych – statystyka opisowa
   1. Dla wybranych danych podstawowe elementy statystyki opisowej
   2. Zapisywanie danych w bazie np. z poszczególnych dni, miesięcy itd.
3. Skrypty testujące hipotezy
   1. Badanie normalności rozkładu, analiza równości wariancji
   2. Testy dla zmiennych zależnych, niezależnych
   3. Testy dla wielu średnich (analiza wariancji)
4. Skrypty wykorzystujące regresję do prognozowania przyszłych wartości
   1. Różne rodzaje regresji – do przewidywania przyszłych wartości
5. Procedury, funkcje, wyzwalacze obsługujące bazę
   1. Dodawanie, usuwanie, aktualizacja rekordów
   2. Archiwizacja usuniętych danych
   3. Logowanie informacji do tabeli
   4. Obsługa wyjątków, również własne wyjątki
   5. Procedury, funkcje z parametrami, możliwe parametry domyślne
   6. Sprawdzanie poprawności dodawanych danych (np. funkcja sprawdzająca poprawność pesela)
6. Procedury, funkcje, wyzwalacze tworzące podsumowania
   1. Zestawienia miesięczne, kwartalne, roczne, w zależności od różnych parametrów – zapisywane w bazie danych (gotowe do wyświetlania na wykresach)
7. Skrypt wizualizujący dane z punktu 7.