**Wymagania dotyczące projektu zaliczeniowego**

Przypominam, że ocena raportu stanowi 50% oceny z ćwiczeń.

Wersja 1- Modelowanie statystyczne

1. Wybrać dane własne lub z jakiegoś repozytorium danych. Nie można dopuścić by dwie osoby robiły to samo opracowanie (gdy zdarzy się taka sytuacja prace nie będą sprawdzane). Aby uniknąć problemów, proszę wpisać wybrany temat w pliku Excel umieszczonym na Teams oraz sprawdzić, czy ktoś nie ma tego samego tematu.
2. Opisać wybrany zestaw danych.
3. Wybrać dwie cechy do opracowania statystycznego.
4. Dla każdej cechy wyznaczyć parametry statystyczne bezpośrednio z próby, zinterpretować wyniki. Porównać wyniki obu cech. Dokonać porównania z rozkładem normalnym (reguła 3-sigma, asymetria itp.)
5. Dla każdej cechy dokonać uporządkowania w szereg rozdzielczy. Wyznaczyć parametry statystyczne, zinterpretować wyniki, porównać z wynikami uzyskanymi bezpośrednio z próby. Porównać wyniki obu cech.
6. Powtórzyć czynności z punktu 5. przy innym uporządkowaniu w szereg rozdzielczy porównać wyniki z tymi uzyskanymi w pkt 5.
7. Dla każdej próby postawić hipotezę zgodności z rozkładem teoretycznym (przetestować dwa rozkłady, w tym koniecznie rozkład normalny). Wykonać testy zgodności (kilka wersji dla różnych wartości alfa) chi-kwadrat i lambda-Kołmogorowa.
8. Wynik testów dla rozkładu normalnego porównać z wynikami z punktu 4.

Elementy oceny:

1. Wybór danych, sformułowanie zadania badawczego.
2. Opis danych
3. Przejrzystość raportu, wykresy, poprawność formalna i językowa
4. Poprawne obliczenia
5. Poprawne budowanie szeregów rozdzielczych
6. Poprawne wnioski
7. Poprawne formułowanie hipotez i wnioskowanie
8. Termin oddania pracy

Wersja 2- Korelacja i regresja

1. Wybrać dane własne lub z jakiegoś repozytorium danych najlepiej takie, dla których jest zadanie typu Regression. Nie można dopuścić by dwie osoby robiły to samo opracowanie (gdy zdarzy się taka sytuacja prace nie będą sprawdzane). Aby uniknąć problemów, proszę wpisać wybrany temat w pliku Excel umieszczonym na Teams oraz sprawdzić, czy ktoś nie ma tego samego tematu.
2. Opisać wybrany zestaw danych.
3. Wybrać co najmniej cztery zmienne objaśniające. Współczynnik zmienności każdej zmiennej powinien być większy od 10%.
4. Sformułować cel badawczy
5. Zbudować model regresji wielorakiej dla wszystkich wybranych zmiennych. Wyznaczyć współczynnik determinacji i zinterpretować jego wartość. Wykorzystać model do prognozy wartości zmiennej objaśnianej.
6. Zastosować metodę Pawłowskiego do redukcji zmiennych (wybrać k=2). Zbudować model regresji dla wybranych zmiennych. Wyznaczyć współczynnik determinacji i zinterpretować jego wartość. Wykorzystać model do prognozy wartości zmiennej objaśnianej.
7. Zastosować metodę analizy współczynników korelacji do redukcji zmiennych.
8. Zbudować model regresji dla wybranych zmiennych. Wyznaczyć współczynnik determinacji i zinterpretować jego wartość. Wykorzystać model do prognozy wartości zmiennej objaśnianej.
9. Dla każdej zmiennej z zestawu otrzymanego w kroku 6. oddzielnie
   1. zastosować analizę korelacji tzn. wyznaczyć prostą regresji i wykonać testy istotności współczynnika korelacji.
   2. wybrać model/modele nieliniowy i wykonać dla niego analizę korelacji
   3. zbudować model wielomianowy
   4. wybrać model najlepiej dopasowany i wykorzystać ten model do prognozy wartości zmiennej objaśnianej

Elementy oceny:

1. Wybór danych, sformułowanie zadania badawczego.
2. Opis danych
3. Przejrzystość raportu, wykresy, poprawność formalna i językowa
4. Poprawne obliczenia
5. Poprawne wnioski dotyczące każdego kroku badań oraz porównania wyników
6. Poprawne formułowanie hipotez w teście istotności współczynnika korelacji i wnioskowanie
7. Termin oddania pracy

Sposób dostarczenia do oceny:

Na platformę Moodle przysłać do terminu podanego na Moodle

1. Plik arkusza kalkulacyjnego zawierający wszystkie dane i wszystkie obliczenia Ewentualnie kod Pythona/ Jupitera, jeśli były wykorzystywane do analizy danych.
2. Pliki z raportem zawierającym tezy, wyniki, wykresy, analizę i podsumowanie w postaci dokumentów pdf.

**Uwagi dotyczące pisania sprawozdań i raportów:**

1. Pamiętamy o tytule sprawozdania i autorze, numeracji stron, nagłówkach tytułów rozdziałów i podrozdziałów, akapitach, jednolitym formatowaniu i stylach.
2. Stosujemy czcionkę 12pkt w tekście, nagłówki 14-16pkt Bold.
3. Tekst justujemy
4. Tytuły i nagłówki trzymamy razem z treścią-- nie mogą zostawać na końcu strony a treść na nowej stronie.
5. Pojedyncze litery (wiszące litery) z końca linii przenosimy na początek następnej (SHIFT+ENTER)
6. Piszemy poprawnie po polsku, bez błędów ortograficznych, stylistycznych, gramatycznych, przestankowych
7. Diagramy i wykresy (inne rysunki) podpisujemy nadając im kolejne liczby jako numery
8. Tabele (np. szeregi rozdzielcze) zamiast podpisu mają nagłówek, w którym umieszczony jest numer i nazwa.
9. Raport ma być przejrzysty, estetyczny i czytelny (nie wklejamy zrzutów ekranów z Excela, tylko piszemy odpowiedni tekst lub tabele i formatujemy)