

Opis funkcjonalności oprogramowania statystycznego dla Uniwersytetu Jana Kochanowskiego w Kielcach

Cechy użytkowe oprogramowania:

- z oprogramowania mogą korzystać przez okres jednego roku **od 1 grudnia 2017r. do 30 listopada 2018r.** do celów edukacyjnych i badawczych wszyscy pracownicy i studenci Uniwersytetu Jana Kochanowskiego w Kielcach;
- możliwość instalacji oraz korzystania z oprogramowania na domowych komputerach pracowników oraz studentów Uniwersytetu Jana Kochanowskiego w Kielcach;
- polskojęzyczne środowisko pracy w programie;
- uaktualnienia do nowych wersji w czasie obowiązywania umowy (bez dodatkowych opłat);
- użytkownicy mają prawo do pomocy technicznej bez wnoszenia dodatkowych opłat (pomoc techniczna jest świadczona za pośrednictwem poczty elektronicznej i telefonicznie w godzinach pracy biura Dostawcy);
- możliwość instalacji zarówno sieciowo, jak i na pojedynczych stanowiskach.

Środowisko pracy z programem i korzystanie z zewnętrznych danych

- Dane mogą być składowane w arkuszu danych umożliwiającym interakcyjne wprowadzanie i przekształcanie danych (sortowanie, transformacje zmiennych, ułoż w stertę/rozrzuc po zmiennych) oraz import i eksport danych (m.in. z plików Excel i plików tekstowych).
- Oprogramowanie ma możliwość łączenia z bazami danych przez OLE DB.
- Wczytywanie i zapis danych w formacie Excel (w wersji 2003 .xls, 2007 xlsx oraz 2010 xlsx), tekstowym, html.
- Wczytywanie i zapis plików danych w formatach: STATISTICA, SPSS, SAS, JMP, Minitab.
- Oprogramowanie zawiera wbudowany, zgodny ze standardami język programowania Visual Basic, który umożliwia dostęp programowy do funkcji programu, programowanie własnych procedur analitycznych (w tym węzłów analizy wykorzystywanych w przestrzeni roboczej data mining) oraz automatyzację prac.
- Dostęp do aplikacji poprzez interfejs COM.
- Oprogramowanie działa na stanowisku komputerowym pod kontrolą systemu operacyjnego Windows Vista/7/8/10 i ich odpowiednikach serwerowych.
- Możliwość instalacji wersji dedykowanej pod 32- lub 64- bitowy system Windows.

Zarządzanie wynikami:

- Oprogramowanie zapewnia możliwość tworzenia raportów z analizy, z możliwością zapisania w formacie PDF.
- Przesyłanie wyników (tabel, wykresów) do dokumentów edytora tekstowego (np. MsWord).
- Możliwość ustawienia wynikowych tabel jako dane wejściowe dla kolejnych analiz.
- Raport otrzymywany przy pomocy oprogramowania przypomina dokument edytora tekstu, a poszczególne obiekty (np. wykresy, arkusze, arkusz czy wykres MS Excel) umieszczone są w nim kolejno, jeden za drugim. Wszystkie raporty mogą być zapisywane nie tylko we własnym formacie oprogramowania, ale także w postaci plików RTF, HTML.
- Oprogramowanie pozwala na zapis wszystkich dokumentów (arkuszy danych i wyników, raporty) w postaci plików HTML, gotowych do opublikowania w Internecie lub Intranecie.
- Możliwość aktualizacji wykresów utworzonych po ich wstawieniu do dokumentu edytora tekstowego (tzn. wykresy mogą być wstawiane jako obiekty OLE).
- Możliwość aktualizacji utworzonych wykresów po zmianie danych źródłowych automatycznie lub ręcznie przez użytkownika (nie dotyczy map).
- Możliwość edycji wykresów po ich utworzeniu.

Funkcjonalność oprogramowania:

Oprogramowanie udostępnia w jednym środowisku użytkownika następujące funkcje analityczne:

- Statystyki podstawowe i tabele
- Możliwość wykonywania analiz w grupach

Załącznik nr 1

- Wykresy: histogramy, wykresy rozrzutu, wykres workowy, wykresy średnia i błędy, wykresy ramkujące, wykres składowych zmienności, wykresy zakresu, wykres rozrzutu z błędem, obrazkowe wykresy rozrzutu, wykresy rozrzutu z rysunkami, wykresy rozrzutu z histogramami, wykresy normalności, wykresy kwantyl-kwantyl, wykresy prawdopodobieństwo-prawdopodobieństwo, wykresy słupkowe/kolumnowe, wykresy liniowe, wykresy sekwencyjne/nakładane, wykresy kołowe, wykresy brakujących danych i spoza zakresu, histogramy dwóch zmiennych, wykresy powierzchniowe, wykresy warstwowe, wykresy wafłowe, wykresy trójkątne, skategoryzowane wykresy XYZ, skategoryzowane wykresy trójkątne, wykresy macierzowe, wykresy obrazkowe, wykresy XYZ 3W, wykresy trójkątne 3W
- Dopasowanie rozkładów
- Regresja wieloraka
- Analiza wariancji (ANOVA)
- Statystyki nieparametryczne
- Rozkłady i symulacje
- Ogólne modele liniowe
- Uogólnione modele liniowe i nieliniowe
- Ogólne modele regresji
- Modele cząstkowych najmniejszych kwadratów
- Komponenty wariancyjne
- Analiza przeżycia
- Estymacja nieliniowa
- Linearyzowana regresja nieliniowa
- Analiza log-liniowa tabel licznosci
- Szeregi czasowe i prognozowanie
- Modelowanie równań strukturalnych
- Analiza skupień
- Analiza czynnikowa
- Składowe główne i klasyfikacja
- Algorytm NIPALS dla analizy składowych głównych i metody cząstkowych najmniejszych kwadratów
- Analiza kanoniczna
- Analiza rzetelności i pozycji
- Drzewa klasyfikacyjne
- Analiza korespondencji
- Skalowanie wielowymiarowe
- Analiza dyskryminacyjna
- Ogólne modele analizy dyskryminacyjnej
- Analiza Mocy Testów
- Sieci neuronowe
- Dobór i eliminacja zmiennych (dla dużych zbiorów danych)
- Analiza koszykowa
- Interakcyjne drążenie danych
- Analiza skupień uogólnioną metodą EM i k-średnich
- Uogólnione modele addytywne (GAM)
- Ogólne modele drzew klasyfikacyjnych i regresyjnych (GTrees)
- Ogólne modele CHAID (Chi-square Automatic Interaction Detection)
- Interakcyjne drzewa klasyfikacyjna i regresyjne
- Wzmacniane drzewa klasyfikacyjne i regresyjne (Boosted Rrees)
- Multivariate Adaptive Regression Splines (MAR Splines)
- Obliczanie dobroci dopasowania
- Szybkie wdrażanie modeli predykcyjnych
- Naiwny klasyfikator Bayesa
- Support Vector Machines
- Metoda k-najbliższych sąsiadów
- Łączenie grup (klas) z wykorzystaniem algorytmu CHAID
- ICA (Independent Component Analysis)

Załącznik nr 1

- Losowy las (Random Forests)
- Standardowe karty kontrolne: karta \bar{X} średniego i R, karta \bar{X} średniego i S, karta pojedynczych obserwacji i ruchomego rozstępu (I/MR), karta sum skumulowanych (CUSUM), karta średniej ruchomej (MA), karta wykładniczo ważonej średniej ruchomej (EWMA), karty dla pomiarów alternatywnych (C, U, Np, P), karta Pareto, karty wielowymiarowe, karty wielotorowe
- Interaktywne zaznaczanie i etykietowanie punktów
- Przypisywanie przyczyn i działań
- Elastyczny, dostosowywalny system alarmowania
- Praca inżyniera i operatora; zabezpieczanie hasłem
- Karty krótkich serii
- Karty wieloźródłowe (zgrupowane i zgrupowane krótkich serii)
- Wskaźniki zdolności, wykonania i linie kontrolne dla rozkładów innych niż normalny
- Karty kontrolne w czasie rzeczywistym; zewnętrzne źródła danych
- Wielowymiarowe karty kontrolne Kart T^2 Hotellinga
- Wielowymiarowe karty kontrolne Wieloźródłowych (zgrupowanych) kart T^2 Hotellinga
- Wielowymiarowe karty kontrolne wykładniczo ważonej średniej ruchomej (MEWMA)
- Wielowymiarowe karty sum skumulowanych (MCUSUM)
- Karta uogólnionej wariancji
- Analiza zdolności procesu: wskaźniki zdolności procesów (np. Cp, Cr, Cpk, Cpl, Cpu, K, Cpm, Pp, Pr, Ppk, Ppl, Ppu i inne),
- Plany badania i analiza powtarzalności i odtwarzalności pomiarów (R&R)
- Analiza Weibulla
- Analiza doświadczenia: Ogólne możliwości
- Analiza resztowa i przekształcenia
- Optymalizacja pojedynczej lub wielu wielkości wyjściowych:
- Standardowe plany frakcyjne dwuwartościowe $2(k-p)$
- Plany frakcyjne $2(k-p)$ o najmniejszej aberracji i maksymalnym niewyikłaniu
- Plany eliminacyjne (Placketta-Burmana)
- Plany frakcyjne trójwartościowe typu $3(k-p)$ z podziałem na bloki oraz plany Boxa-Behnkena
- Plany centralne kompozycyjne (powierzchnia odpowiedzi)
- Plany kwadratów łańcuchowych
- Doświadczenia wg metody Taguchi
- Plany dla mieszanin i powierzchni o podstawie trójkątnej
- Plany dla ograniczonych powierzchni i mieszanin
- Plany D i A-optymalne
- Funkcjonalność text mining
- Analiza dokumentów zapisanych w formacie MS Word
- Zliczanie wystąpień słów
- Różne miary częstości występowania słów : prosta częstość, częstość binarna (ang. binary frequency), odwrotna częstość dokumentowa (ang. inverse document frequency), częstość logarytmiczna
- Możliwość określania własnej stop-listy
- Możliwość określania synonimów
- Wykonywanie rozkładu według wartości osobliwych (ang. singular value decomposition) dla miar częstości występowania słów w zbiorze dokumentów
- Analiza podstawowych przyczyn
- Optymalizacja wielkości wyjściowych
- Ogólna optymalizacja
- Wdrażanie modelu MSPC
- Analiza składowych głównych (PCA)
- Częstkowe najmniejsze kwadraty (PLS)
- Wielokierunkowe częstkowe najmniejsze kwadraty wg partii (BMPLS)
- Wielokierunkowa analiza składowych głównych według czasu (TMPCA)
- Wielokierunkowe częstkowe najmniejsze kwadraty wg czasu (TMPLS)
- Wykrywanie reguł asocjacji

Załącznik nr 1

- Analiza sekwencji
- Analiza skojarzeń
- Wykresy zmienności,
- Wykresy wielokrotne, pozwalające bezpośrednio porównywać wiele zmiennych zależnych,
- Komponenty wariancyjne z przedziałami ufności,
- Elastyczne operowanie wieloma zmiennymi zależnymi: jednoczesne analizowanie wielu zmiennych wg tego samego lub różnych planów,
- Wykresy komponentów wariancyjnych
- Tabele raportujące
- Definiowanie reguł poprawności danych
- Reguły poprawności danych
- Analiza brakujących danych
- Przekodowanie na zmienne sztuczne
- Szybkie rekodowanie
- Przekształcenia zmiennych
- Zliczanie wystąpień
- Porządkuj zmienne wielokrotnych odpowiedzi
- Kalkulator liczebności próby
- Wazenie wieńcowe przypadków
- Propensity score matching
- Podsumowanie skali pozycyjnej
- Podsumowanie skali rangowej
- Wykres dyferencjału semantycznego
- Wykres dla skali Stapela
- Rzetelność skali
- Metoda ocen porównawczych Thurstone'a
- Współczynniki zgodności sędziów
- Krzywe ROC
- Metaanaliza i metaregresja
- Kreator regresji logistycznej
- Kreator regresji liniowej
- Analiza conjoint
- Aglomeracja z punktem odcięcia
- Analiza PROFIT
- Uogólniona metoda składowych głównych (PCA)
- Porządkowanie liniowe
- Miary powiązania/efektów dla tabel 2x2
- Analiza koncentracji
- Standaryzowane miary efektu
- Test post hoc ANOVA Friedmana
- CATANOVA
- Karta CUSUM ważona ryzykiem
- Indeks KMO oraz Test sferyczności Bartletta
- Wykres Blanda-Altmana
- Wykres słupkowy (kolorowe słupki)
- Wykres sekwencyjny
- Wykres radarowy
- Wykres mozaikowy
- Wykres kołowy (SPie plot)
- Możliwość wizualizacji danych na mapach:

Załącznik nr 1

- a) Gotowe szablony map dostępne w programie obejmują podział Polski na: województwa, powiaty, gminy, okręgi wyborcze, województwa w podziale na powiaty, województwa w podziale na gminy, województwa w starym podziale
- b) Możliwość wczytywania innych niż zawarte w programie szablonów map w formacie *.shp
- c) Kolorowanie na mapach tła obszarów wartościami zadanej zmiennej (predefiniowane palety do wyboru, możliwość ustalenia palety użytkownika, możliwość ustalenia własnych granic dla przedziałów legendy, możliwość zapisu/wczytania palety kolorów z/do pliku)
- d) Generowanie wykresów kołowych i słupkowych (możliwość ręcznej zmiany wielkości wykresu, możliwość ręcznego ustalenia jego położenia, możliwość zmiany skalowania wysokości słupka względem wiersza/kolumny/całości, zmienny promień wykresu kołowego zależny od wartości ze zmiennej)
- e) Wyświetlanie etykiet tekstowych pobranych z zadanej zmiennej lub zmiennej zawierającej mapowanie elementów wraz z formatowaniem zadanych przez użytkownika (kolor, krój itp.), oraz ręczną korekcją położenia etykiety względem innych elementów wykresu
- f) Różne stany wyświetlania elementów obszaru – aktywny, nieaktywny, ukryty
- g) Rodzaj i grubość linii rysowanych jako granice może być zmieniana przez użytkownika
- h) Możliwość zapisu/odczytu z i do pliku wszystkich opcji wyglądu mapy
- i) Możliwość ręcznej edycji przez użytkownika szablonów map wczytanych w programie (usuwanie obszarów, scalanie obszarów) i zapisu jako nowy szablon
- j) Możliwość zarejestrowania wygenerowanej mapy (z wizualizacją danych) w postaci makra