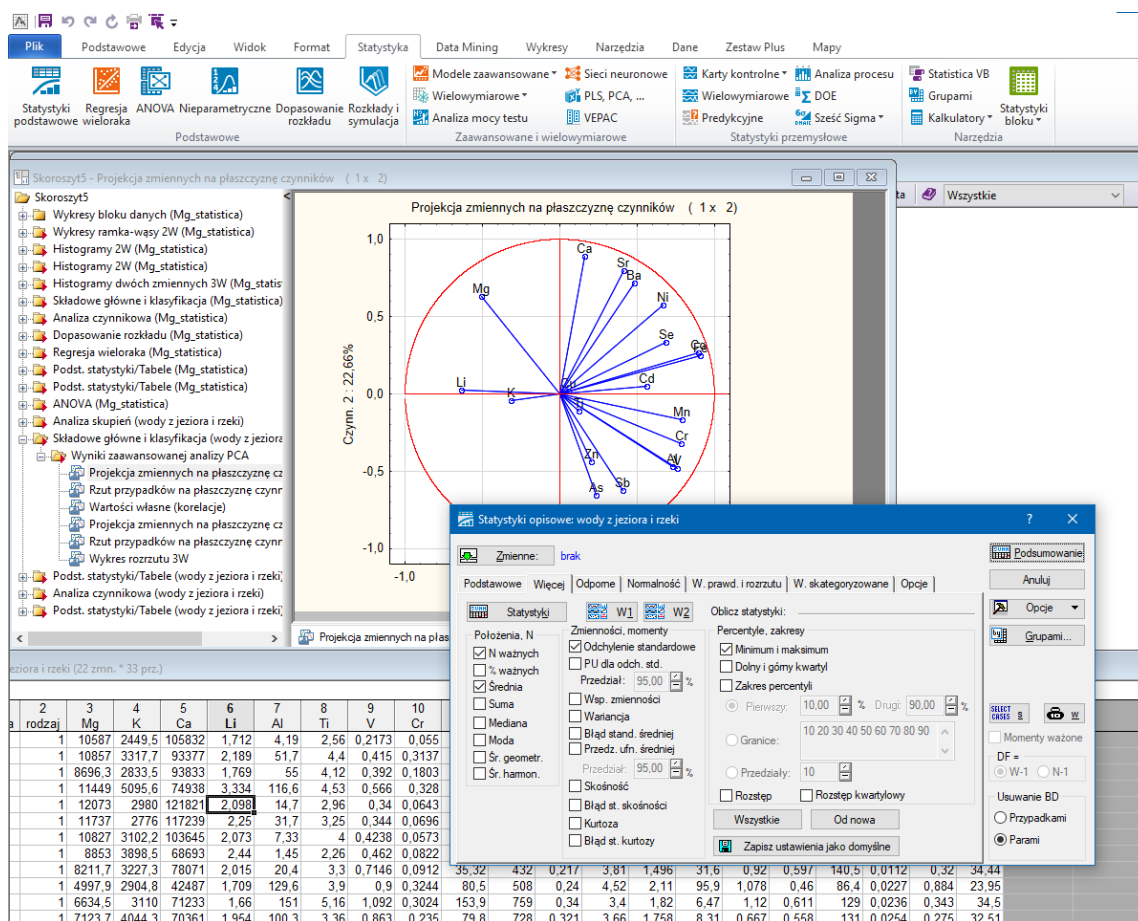


Ćwiczenie 2K2

STATYSTYCZNA OCENA I WIZUALIZACJA WYNIKÓW OZNACZANIA Mg W ŻYWNOSCI STAŁEJ Z UŻYCIEM PROGRAMU STATISTICA

Statistica jest zaawansowanym programem do analizy danych, stosowanym przez profesjonalistów z wielu dziedzin, m.in. nauk ścisłych, ekonomii i medycyny. Pozwala na przeprowadzenie w łatwy sposób testów statystycznych, sporządzenie wykresów, obliczenie podstawowych parametrów statystyki opisowej, jak również wykonanie zaawansowanych testów i analizy wielowymiarowe. Użytkownik nie musi wprowadzać ręcznie wzorów matematycznych i nie jest wymagana znajomość formuł tekstowych, tak jak w przypadku programu Excel. Wszystkie operacje można przeprowadzić po kliknięciu w odpowiednią ikonę na wstążce u góry okna programu.



Rys. 1 Przykładowe okno programu Statistica. U góry wstążka z zakładkami i ikonkami; po lewej foldery z wykresami; na środku wykres – projekcja składowych głównych (PCA); na dole arkusz z danymi; po prawej okno dialogowe podstawowych parametrów statystycznych.

Cel ćwiczenia

Celem ćwiczenia jest zapoznanie się z obsługą i podstawowymi funkcjami programu Statistica, wykonanie popularnych testów statystycznych i wykresów w celu oceny danych pomiarowych, a także przeprowadzenie analizy chemometrycznej na przykładzie danych wielowymiarowych.

Wykonanie ćwiczenia

Wprowadzenie danych

- w Excelu, w nowym arkuszu ustaw dane pomiarowe z poprzedniego ćwiczenia (zawartość Fe w suchej masie mleka w proszku)

- stwórz tabelę z 4 kolumnami;
- nagłówki: *nazwa próbki*, *rodzaj próbki*, *Fe AAS*, *Fe ICP*

Jako rodzaj próbki wpisz numer; każdy analizowany produkt ma kolejny numer (1, 2, 3).

- w Statistica, wklej dane do nowego arkusza; ustaw nazwy nagłówków kolumn, zaznacz nagłówki kolumny z zawartościami Fe → opcja „Specyfikacja zmiennej” i ustaw liczbę miejsc dziesiętnych na 1.

W arkuszu z danymi kolumny (analizy) nazywają się **zmiennymi**, a wiersze (próbki) – **przypadkami**.

- usuń puste zmienne w arkuszu, jeśli występują.

Wizualizacja wyników i podstawowe funkcje Statistica

- Zaznacz dane pomiarowe, klikając na nagłówki kolumny ze stężeniami i w opcjach „Statystyki bloku danych” i „Wykresy danych wejściowych” wybierz odpowiednie narzędzie aby wygenerować:

- **Statystyki podstawowe** dla kolumn (Wszystkie).
 - **Histogram** (dopasuj normalny)
 - zmień zakres osi x → dwuklik na wykres → opcja Histogram → zaznacz Granice → Określ granice: minimum, maksimum, krok
 - zmień tytuł osi y na „Liczba pomiarów” → dwuklik na tytuł osi y lub prawy klik → Tytuł
 - rozjaśnij wykres z pomocą suwaka pod wykresem
 - najedź kursorem na oś x lub y, gdy cursor zmieni wygląd na łapkę przesuń trzymając lewy przycisk myszki; W zależności, czy cursor znajduje się na krawędzi czy na środku osi łapka będzie miała inny wygląd i będzie, odpowiednio, zmieniać zakres osi lub przesuwać wykres.
 - **Histogram 2W** względem: X – zmierzone stężenia, Kategorie – rodzaj próbki.
 - **Wykres ramka-wąsy** → Mediana/kwartyle/rozstęp.
 - umieść histogramy 2W i wykres ramka-wąsy w jednym folderze
 - Powiększ dwukrotnie wykresy i umieść je w dwóch kolumnach:
prawy klik na folder → Podgląd folderu → Ustawienia (+ zastosuj we wszystkich folderach).
- Klikając na folder uzyskuje się podgląd na wszystkie znajdujące się w nim wykresy.

Drugim sposobem tworzenia wykresów jest kliknięcie na ikonkę we wstążce na górze okna programu.

W zakładce „Wykresy” znajdują się popularne wykresy. Klikając na ikonki wykresów otwierają się dodatkowe opcje konfiguracji, których zabrakło w zadaniu poprzednim.

- wybierz wykres „**Histogram**”,

- zaznacz „Zmienne” (wybierz stężenia), zaznacz „Pokaż procenty”, „Test Shapiro Wilka”, „Statystyki opisowe”, „Całkowita licznosc”;

- ustaw pole tekstowe z wynikami testów na wykresie tak, aby nie zasłaniało żadnej osi ani danych na wykresie (ustaw skalę na osiach)

- wybierz wykres „**Ramka-wąsy**” → Rodzaj wykresu „Ramka-wąsy” i „Wielokrotny”. Wybierz Zmienne: zmienną zależną są zawsze wyniki pomiarów w jednostce stężenia/zawartości czyli wartości ciągłe, a zmienną grupującą jest zmienna dyskretna, czyli jakościowa, a w naszym przykładzie będzie to *rodzaj próbki*.

- utwórz wykres ramka-wąsy względem zmiennej grupującej *rodzaj próbki*; w zakładce „Więcej” znajdują się wszystkie opcje wykresu: zaznacz „Pokaż dane surowe”, „Połącz punkty środkowe” i „Separacja punktów, losowa”; zaznacz „Statystyki”, zaznacz „Test F i p (ANOVA)”.

Aby szybko utworzyć drugi wykres, wychodząc z tych samych ustawień, kliknij w pasek „Wykres ramka...” na dole okna.

- ustaw wielkość znaczników „Dane surowe” na 3, a „Mediana” na 6 → dwuklik na znacznik → przycisk „Znaczniki”.

- dodaj znacznik „Średnia” → dwuklik na tło wykresu → W opcjach „Wykres właściwy / Ogólne” wybierz z listy na górze „5. Średnia” i „Wyświetl wykres”.

- ustaw pole tekstowe z wynikami testów na wykresie tak, aby nie zasłaniało żadnej osi ani danych na wykresie (ustaw skalę na osiach).

Testy statystyczne i statystyka opisowa

W zakładce „Statystyka” znajdują się wszystkie podstawowe i zaawansowane narzędzia do oceny i prezentacji wyników.

- Oblicz podstawowe parametry statystyczne wybierając ikonkę „**Statystyki podstawowe**” i dalej „Statystyki opisowe”; Aby zatwierdzić test w oknie dialogowym klikaj w przycisk „Podsumowanie”:

- w zakładce „Więcej” zaznacz „Suma”, „Mediana”, „Moda”, „Skośność”, „Kurtoza”, „Dolny i górny kwartył” i „Rozstęp”.

- wykonaj „test t dla grup niezależnych, względem zmiennych”, jako zmienne wybierz po jednej grupie wyników pomiarów dla każdej listy zmiennych (grup), kliknij przycisk „Grupami”, „Zmienna grupująca” i wybierz *rodzaj próbki*, zaznacz „Wyniki w jednym folderze” i „Zbierz wyniki w jednej tabeli”.

- wykonaj test ANOVA → „Jednoczynnikowa ANOVA”, test wykonaj osobno dla każdej zmiennej zależnej, czyli stężenia analitu ; zaznacz „Średnie/wykresy” i „Wszystkie efekty”; Wyświetl wszystkie wyniki testu klikając na folder.

Analiza wielowymiarowa

Analiza wielowymiarowa zostanie przeprowadzona dla danych wielowymiarowych, tzn. że dla jednej próbki zmierzona więcej niż jeden parametr, np. stężenia kilku pierwiastków, ale mogą to być dowolne parametry pomiarowe, takie jak: barwa, temperatura, CHZT, przewodność, chlorki, tlen rozpuszczony, azot całkowity,

krzemionka itd. Analiza wielowymiarowa pozwala na wskazanie zależności między parametrami, które nie są widoczne gołym okiem. Drugim zestawem danych wielowymiarowych są wyniki pomiaru pierwiastków w ścianach naczyń krwionośnych pobranych od pacjentów chorujących na zaawansowaną miażdżycę.

- ściągnij ze strony plik z danymi; sprawdź w excelu jak wyglądają dane, czy nie ma brakujących danych, czy nagłówki są prawidłowo opisane

- w Statistice: zakładka „Podstawowe” → „Nowy” → „Przestrzeń robocza”

- wybierz „Pusta” i OK

- Gdy program zapyta o źródło danych → „Pliki”; w rodzaju plików wybierz „Wszystkie pliki *.*” i zaznacz ściągnięty plik Excel z danymi → „Import wszystkich arkuszy” → zaznacz „Weź nazwy zmiennych...” i „Weź nazwy przypadków ...”

- Gdy pojawi się nowa ikonka w przestrzeni roboczej kliknij ją aby otworzyć arkusz z danymi

- utwórz **macierz korelacji** w „Statystyki podstawowe” → „Jedna lista zmiennych”, zaznacz wszystkie pierwiastki → „Podsumowanie”.

- wykonaj **analizę składowych głównych (PCA)** → zakładka „Statystyka” → „Wielowymiarowe”.

- w zmiennych do analizy wybierz wszystkie pierwiastki, a jako zmienną grupującą wybierz *rodzaj*; Pamiętaj, że w pierwszej kolumnie numerami oznaczono rodzaje próbek.

- Wygeneruj „Własności własne”, „wykres osypiska” i na jego podstawie przeanalizuj udział procentowy pierwszych składowych w całkowitej zmienności danych

- Wygeneruj „Wykres 2W współrz. czyn. zmn.” i „Wykres 2W współrz. czyn. przyp.”

Wymagania

- sposoby wizualizacji danych (histogram, wykres skrzynka-wąsy);
- ANOVA – definicja, zastosowanie
- Analiza składowych głównych – cel, zastosowanie
- Definicje i zastosowanie: analiza wielowymiarowa, chemometria

Literatura

1. „Ocena i kontrola jakości wyników analitycznych”, Wydawnictwo Naukowo-Techniczne Warszawa 2007, praca zbiorowa pod redakcją Piotra Konieczki i Jacka Namieśnika.
2. „Analiza statystyczna w laboratorium analitycznym”, Wojciech Hyk, Zbigniew Stojek;
3. „Podstawy chemii analitycznej tom 1”, D.A. Skoog, D.M. West, F.J. Holler, S.R. Crouch, Wydawnictwo Naukowe PWN, Warszawa 2007, rozdział 7;
4. „Statystyka i chemometria w chemii analitycznej”, J. Miller, przekład J. Mazerski, Wydawnictwo Naukowe PWN, Warszawa 2016
5. <https://www.statsoft.pl/textbook/stathome.html>
6. <https://mirosławmamczur.pl/category/wykresy/page/2/>
7. https://brain.fuw.edu.pl/edu/index.php/Wnioskowanie_Statystyczne_-_wyk%C5%82ad