

2017

# Riki tiki statistiki

DOKUMENTACJA  
PRZEMYSŁAW STODOLNY  
PAWEŁ PAŃCZYK  
DAMIAN PUZIŃ  
DAMIAN KWIATKOWSKI

## Spis treści

1.	Charakterystyka produktu.....	3
1.1	Zastosowanie .....	3
1.2	Cechy wyróżniające .....	3
1.3	Cel produktu .....	4
1.4	Opis użytkowników systemu i ich funkcje .....	4
1.5	Struktura funkcjonalna (hierarchia funkcji).....	4
1.5	Model rozwiązań (główni użytkownicy, sposób funkcjonowania, struktura sprzętowa i jej rozmieszczenie w przestrzeni).....	5
2.	Harmonogram realizacji projektu.....	5
3.	Diagramy UML .....	5
3.1	Przypadki użycia .....	5
3.2	Diagramy sekwencji.....	8

## 1. Charakterystyka produktu

Program statystyczny RIKI TIKI STATISTIKI jest aplikacją służącą do statystycznej analizy danych, tworzenia wykresów, pracy na bazach danych, wykonywania transformacji danych. Zadaniem (może funkcją?) programu jest przetworzenie oraz filtrowanie dużej ilości danych używanych w obliczeniach statystycznych. Następnie wykonanie obliczeń statystycznych oraz prezentacja wyników w postaci graficznej.

System RIKI TIKI STATISTIKI oferuje wybrane operacje statystyczne na danych zawartych w programie Microsoft Excel. Jest skierowany dla studentów, którzy potrzebują prostą w obsłudze aplikację liczącą wybrane funkcje statystyczne. Program charakteryzuje się prostym interfejsem użytkownika. W tabeli nr 1 przedstawione są funkcjonalności programu w porównaniu z konkurencyjnymi programami statystycznymi. Są to Excel oraz Statistica. Porównania dokonano na podstawie podręcznika dla programu Statistica<sup>1</sup>, kompendium aplikacji Excel<sup>2</sup> oraz własnych doświadczeń.

Tabela 1. Porównanie programu RIKI TIKI STATISTIKI z programami konkurencyjnymi

Funkcje	RIKI TIKI STATISTIKI	Statistica	Excel
Mediana	●	●	●
Odchylenie standardowe	●	●	●
Kowariancja	●	●	●
Rozkład istotności	●	●	●
Rozkład prawdopodobieństwa	●	●	●
Łatwość obsługi	●	●	●
Cena	●	●	●
Obciążenie systemu	●	●	●

Rozwiązanie RIKI TIKI STATISTIKI dostępne jest w postaci aplikacji desktopowej. Wspierane systemy operacyjne to Windows XP x32 Sp3, Windows XP x64 Sp2, Windows Vista x32, Windows Vista x64, Windows 7 x32, Windows 7 x64, Windows 8 x32, Windows 8 x64, Windows 8.1 x32, Windows 8.1 x64, Windows 10 x32, Windows 10 x64.

### 1.1 Zastosowanie

Program służy do obliczeń statystycznych. W programie znajdują się funkcje tj.: rozkład prawdopodobieństwa, mediana, odchylenie standardowe, kowariancja, rozkład istotności.

### 1.2 Cechy wyróżniające

- Przejrzysty i zrozumiały interfejs użytkownika
- Program nie wymaga instalacji
- Program nie zajmuje wiele miejsca na dysku

<sup>1</sup> M. Rabiej, *Statystyka z programem Statistica*, Gliwice: Wydawnictwo Helion 2012

<sup>2</sup> J. Walkenbach, *Microsoft Excel 2016 Bible*, Indianapolis: Wydawnictwo John Wiley & Sons, Inc. 2015

### 1.3 Cel produktu

Program pobiera dane z arkusza kalkulacyjnego Microsoft Office. Następnie filtruje dane według określonych kryteriów. Są one automatycznie dobierane przez aplikację po wybraniu przez użytkownika żądanej funkcji statystycznej. Program wykonuje obliczenia statystyczne: rozkład prawdopodobieństwa, medianę, odchylenie standardowe, kowariancję, rozkład istotności. Po obliczeniach aplikacja prezentuje wyniki w postaci liczbowej oraz graficznej oraz eksportuje je do pliku „pdf”.

### 1.4 Opis użytkowników systemu i ich funkcje

- użytkownik standardowy – obsługa programu w zakresie głównych funkcji tj.: importu danych, obliczeń statystycznych wraz z prezentacją wyników oraz eksportu do pliku pdf.

### 1.5 Struktura funkcjonalna (hierarchia funkcji)

Na poniższych rysunkach przedstawiono schemat struktury funkcjonowania (hierarchie funkcji) oprogramowania RIKI TIKI STATISTIKI dla poszczególnych funkcji statystycznych.



Rys. 1. Schemat struktury funkcjonowania dla mediany



Rys. 2. Schemat struktury funkcjonowania dla rozkładu Istotności



Rys. 3. Schemat struktury funkcjonowania dla odchylenia standardowego



Rys. 4. Schemat struktury funkcjonowania dla kowariancji



Rys. 5. Schemat struktury funkcjonowania dla rozkładu prawdopodobieństwa

### 1.5 Model rozwiązań (główni użytkownicy, sposób funkcjonowania, struktura sprzętowa i jej rozmieszczenie w przestrzeni)

- System operacyjny Microsoft Windows
- Arkusz kalkulacyjny Microsoft Excel
- Python 2.7
- Pycharm
- Qt

## 2. Harmonogram realizacji projektu

- 05-11/03/2017 - Utworzenie wstępnej dokumentacji projektu. Podział obowiązków,
- 12-18/03/2017 - Opracowanie funkcji matematycznych dla projektu,
- 19-25/03/2017 - Opracowanie UML - przypadki użycia, diagramy sekwencji. Oddanie dokumentacji,
- 26/03/2017-01/04/2017 - Stworzenie interfejsu użytkownika aplikacji,
- 02-22/04/2017 - Implementacja funkcji matematycznych w kodzie. Import/eksport,
- 22/04/2017 - 13/05/2017 - Testy aplikacji,
- 14/05/2017 - 10/06/2017 - Oddanie gotowej aplikacji do użytku.

## 3. Diagramy UML

### 3.1 Przypadki użycia

Rysunek nr 6 przedstawia diagram przypadków użycia w systemie.

**Import danych z excela** - po wybraniu opcji importu danych wybieramy ścieżkę do pliku. Następnie importujemy dane co spowoduje wyświetlenie danych w głównym oknie.

**Funkcje matematyczne** - wyświetla pięć funkcji matematycznych i pozwala na wybór tej która nas interesuje.

**Rozkład prawdopodobieństwa** - po wybraniu funkcji wybieramy zakres danych, następnie funkcja wykonuje obliczenia przekazuje wyniki do klasy wyniki.

**Rozkład istotność** - po wybraniu funkcji wybieramy zakres danych, następnie funkcja wykonuje obliczenia przekazuje wyniki do klasy wyniki. Funkcja wykonuje test t-studenta dla dwóch populacji. Wynik, wraz z poziomem istotności ( $\alpha=0.05$ ) jest porównywany z tabelą rozkładu t-studenta i na tej podstawie jest stwierdzana istotność.

Test jest wykonywany na podstawie wzoru:

$$T = \frac{x_1 - x_2}{S_{x_1 - x_2}}, \text{ gdzie:}$$

$x_1, x_2$  - średnie arytmetyczne populacji

$$S_{x_1 - x_2} = \sqrt{\frac{(n_1 - 1) * s_1 + (n_2 - 1) * s_2}{n_1 + n_2 - 2} * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}, \text{ gdzie:}$$

$n_1, n_2$  - liczebność populacji 1 i 2;

$s_1, s_2$  - wariancja pierwszej i drugiej populacji;

**Mediana** - po wybraniu funkcji wybieramy zakres danych, następnie funkcja wykonuje obliczenia przekazuje wyniki do klasy wyniki.

Obliczenia są wykonywane zgodnie ze wzorem:

$$Me = \frac{(n+1)}{2}, \text{ gdzie:}$$

$n$  - ilość elementów;

**Kowariancja** - po wybraniu funkcji wybieramy zakres danych, następnie funkcja wykonuje obliczenia przekazuje wyniki do klasy wyniki.

Program liczy kowariancję na podstawie wzoru:

$$\text{cov}(X, Y) = E(X * Y) - (E(X) * E(Y)), \text{ gdzie:}$$

$E(X), E(Y)$  - wartości oczekiwane populacji  $X$  i  $Y$ .

**Odchylenie standardowe** - po wybraniu funkcji wybieramy zakres danych, następnie funkcja wykonuje obliczenia przekazuje wyniki do klasy wyniki.

W pierwszej kolejności program liczy wartość oczekiwaną populacji na podstawie wzoru:

$$\sum_{i=1}^n x_i * p_i, \text{ gdzie:}$$

$x_i$  -  $i$ -ty element populacji;

$p_i$  - prawdopodobieństwo wystąpienia elementu

Następnie, odchylenie standardowe populacji jest liczone zgodnie ze wzorem:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (xi - \mu)^2}{n}} \quad , \text{ gdzie:}$$

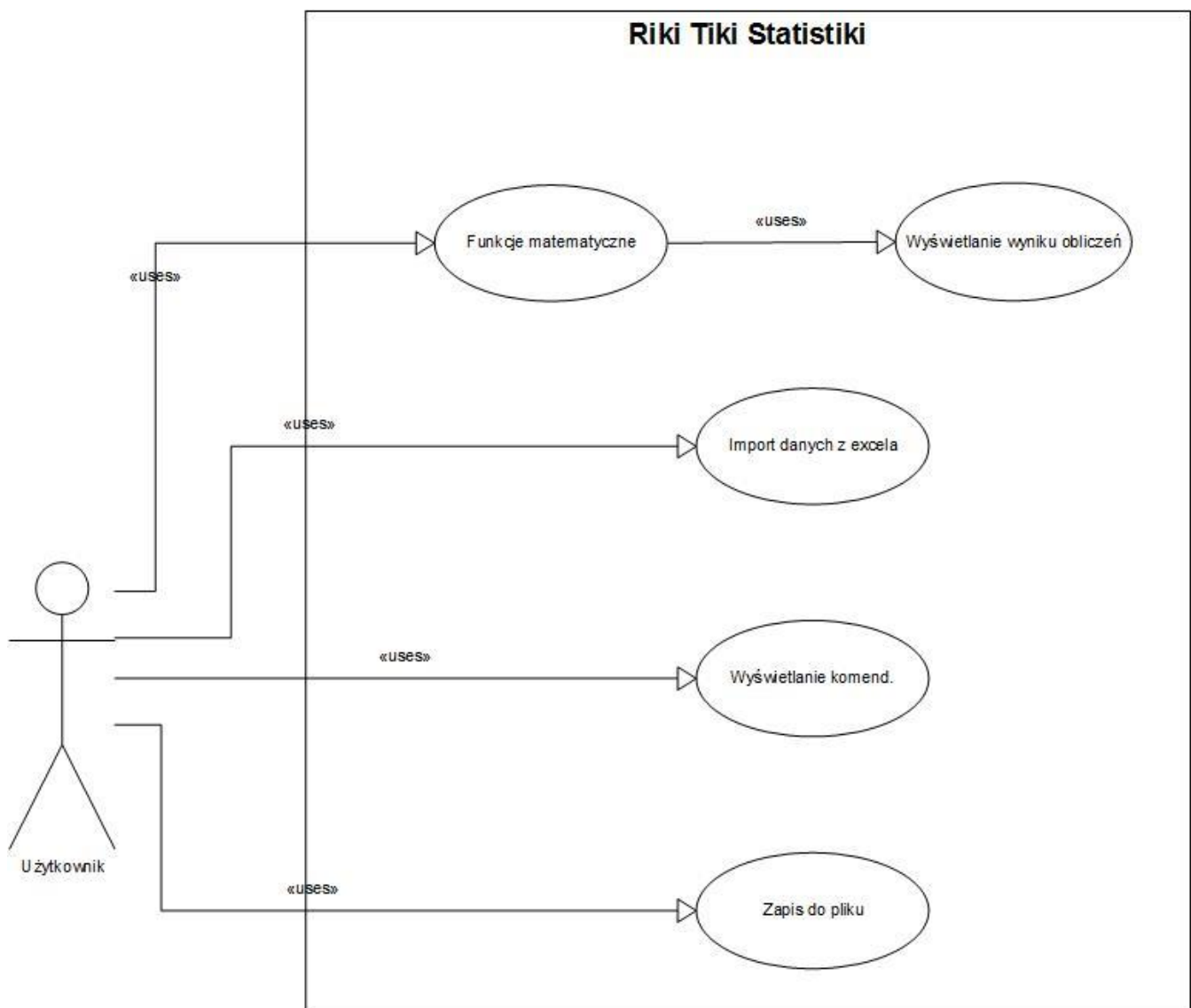
$\sigma$  - odchylenie standardowe;

$n$  - liczba obserwacji/elementów populacji;

$xi$  -  $i$ -ty element;

$\mu$  - wartość oczekiwana;

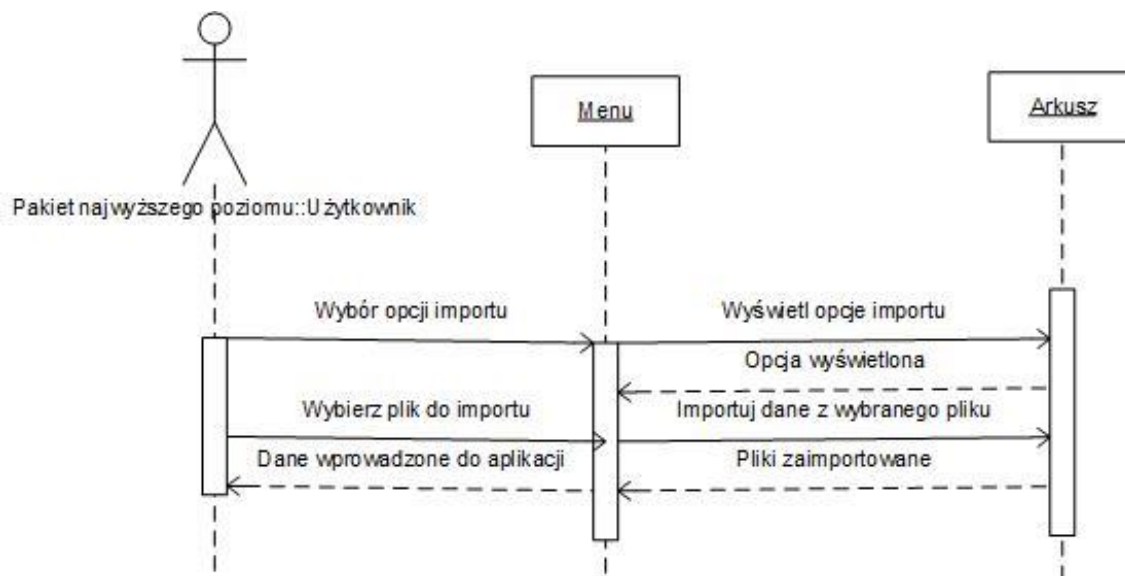
**Zapis do pliku** - wybranie opcji zapisu spowoduje zapis do pliku .pdf wyników obliczeń.



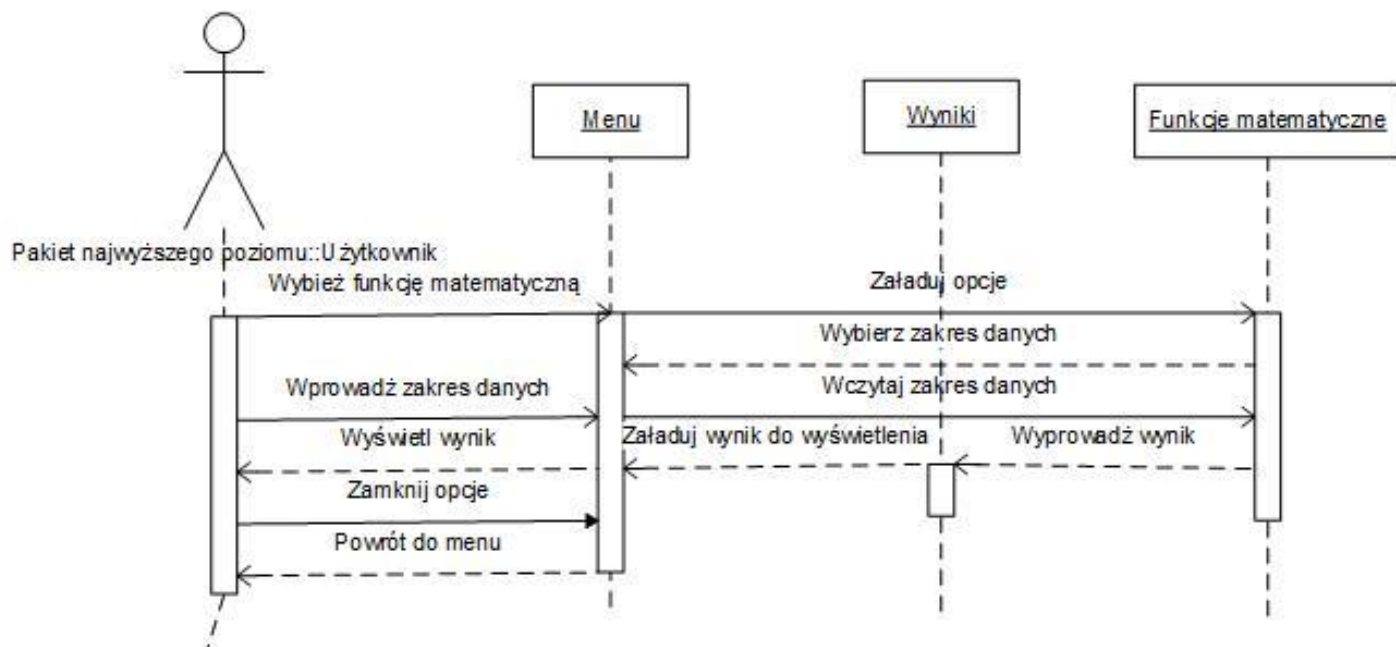
Rys. 6. Przypadki użycia

### 3.2 Diagramy sekwencji

Poniższe diagramy sekwencji przedstawiają przepływ komunikatów pomiędzy danymi obiektami systemu.

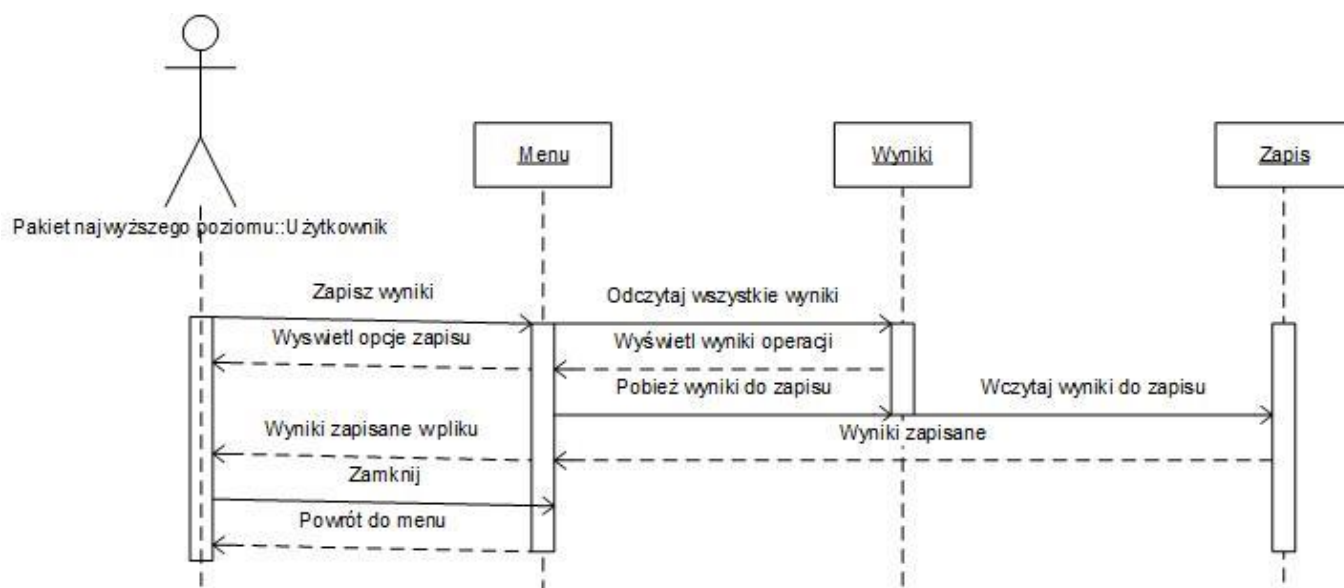


Rys. 7. Import danych bezpośrednio z pliku aplikacji Microsoft Excel. Kolejność wysyłania komunikatów pomiędzy obiektami w celu wykonania importu danych z pliku do aplikacji.



Rys. 8. Funkcje matematyczne. Kolejność wysyłania komunikatów pomiędzy obiektami w celu wykonania wybranej funkcji matematycznej.





Rys. 9. Zapis do pliku „.pdf”. Kolejność wysyłania komunikatów pomiędzy obiektami w celu zapisu wyników funkcji matematycznych do plików „.pdf”.