

## Wstęp

W procesie dydaktycznym, obecnie na każdym poziomie nauczania, codziennością prawie staje się korzystanie z różnego rodzaju narzędzi informatycznych, począwszy od prostych programów, kończąc na zaawansowanych aplikacjach. Mają one ułatwić, a czasami usprawnić i przyspieszyć, proces przetwarzania informacji, proces nauczania i wreszcie proces badawczy. Do takich narzędzi czy systemów wspomagających należy system Stata.

Stata jest jednym z wielu funkcjonujących systemów, pakietów komputerowych, które mogą być wykorzystywane w procesie dydaktycznym wyższej uczelni, w tym w procesie nauczania takiego przedmiotu, jakim jest statystyka.

Podręcznik *Statystyka od podstaw z systemem Stata* został przygotowany jako bliźniaczy podręcznik do podręcznika *Statystyka od podstaw z systemem SAS*, E. Frątczak i A. Korczyńskiego (SGH, Warszawa 2013). Filozofia konstrukcji podręcznika i zbiory wykorzystane w nim do ilustracji przykładów są takie same, jak w podręczniku *Statystyka od podstaw z systemem SAS*. Chodziło nam o to, aby Czytelnicy mogli porównać dwa systemy statystyczne, popularne w środowisku naukowym i studenckim w zakresie podstawowych analiz statystycznych.

Inspiracją do przygotowania podręczników były programy nauczania statystyki, funkcjonujące na uczelniach ekonomicznych, w tym w Szkole Głównej Handlowej (SGH). Wiodącym podręcznikiem do nauczania statystyki według standardowego programu w SGH (aktualny sylabus programu nauczania wraz z obowiązującą literaturą zamieszcza załącznik 3) jest *Statystyka od podstaw*, J. Joźwiak i J. Podgórskiego (2012).

Zakładamy, że informacje ogólne o systemie Stata oraz przykłady zawarte w tym podręczniku wraz z algorytmami rozwiązań mogą być pomocne w rozwiązywaniu ćwiczeń ze statystyki w programie podstawowym statystyki, wspomagany systemem Stata. Zdajemy sobie sprawę, że tekst tego podręcznika prezentuje zaledwie mały wycinek możliwości obu wykorzystanych narzędzi w zakresie "Statystyki od podstaw". Mamy nadzieję, że ocena tego podręcznika i wszelkie uwagi Czytelników pozwolą nam w przyszłości przygotować jego ulepszoną i poszerzoną wersję.

Uważamy, że wykorzystanie tak bogatych narzędzi statystycznych, jakimi są Stata i SAS w procesie dydaktycznym, powinno się zacząć od zapoznania się z podstawami statystyki, a następnie przechodzić do coraz bardziej zaawansowanych procedur, algorytmów i analiz, które mogą stanowić przedmiot wykładów specjalistycznych.

Logika konstrukcji tego podręcznika jest taka, że nie omawiamy procedur statystycznych ani podstaw teorii statystyki, ponieważ Czytelnik powinien je znaleźć w wielu podręcznikach ze statystyki, w tym m.in. we wspomnianych już opracowaniach J. Joźwiak i J. Podgórskiego (2012) oraz E. Frątczak i A. Korczyńskiego (2013).

Struktura podręcznika jest następująca: najpierw omówiono narzędzie wraz z podstawowymi informacjami o możliwościach jego wykorzystania do obróbki i wizualizacji danych wraz z raportowaniem, a następnie, w kolejnych rozdziałach, podano przykłady analizy, począwszy od metody reprezentacyjnej i wybranych schematów losowania próby, skończywszy zaś na analizie szeregów czasowych.

## Opis

Podręcznik *Statystyka od podstaw z systemem Stata* jest podręcznikiem akademickim, łączącym tradycję z nowoczesnością. Tradycję stanowią metody statystyczne, których zastosowania są prezentowane w pracy - nowoczesność to nowa technologia związana z wykorzystaniem systemu Stata. Konwencja pracy wpisuje się w nurt tego, co stanowi zmianę statystycznego myślenia w procesie nauczania statystyki. Statystykę coraz częściej, w nowoczesnym podejściu definiuje się jako: "sposób pozyskiwania informacji z danych" (*Statistics is a way to get information from data*, G. Keller, 2012).

W pracy Czytelnik znajdzie podstawowe informacje o systemie Stata, które ułatwią zrozumienie zarówno istoty metod statystycznych jak i korzystanie z nich. Intencją autorów było położenie głównego nacisku nie na teorię, ale na praktyczne zastosowanie metod statystycznych do różnych zbiorów danych, opis postępowania analitycznego dla wykorzystanych metod wraz z wnioskowaniem na tej podstawie.

Podstawowe zagadnienia omawiane w pracy to: *elementy metody reprezentacyjnej, analiza struktury, rozkłady teoretyczne zmiennych losowych, wnioskowanie statystyczne, analiza wariacji i kowariancji, analiza korelacji i regresji oraz analiza szeregów czasowych*. Wykorzystane w pracy zbiory danych dotyczą zagadnień procesów społecznych, ekonomicznych, biznesowych. Podręcznik jest polecany nie tylko dla studentów uczelni i kierunków ekonomicznych, społecznych, ale również dla analityków biznesowych stawiających pierwsze kroki w analizach statystycznych. Czytelnikom polecamy również podręcznik z tej serii *Statystyka od podstaw z systemem SAS*.

## Spis treści

### Wprowadzenie

#### 1. Środowisko pracy w systemie Stata

##### 1.1. Informacje ogólne

##### 1.2. Interfejs pakietu Stata

###### 1.2.1. Okna

###### 1.2.2. Pasek narzędzi

###### 1.2.3. Pasek statusu

##### 1.3. Składnia i sposób wpisywania komend

##### 1.4. Pomoc i źródła informacji o systemie Stata

##### 1.5. Aktualizacja pakietu Stata

## **2. Przetwarzanie danych w systemie Stata**

### 2.1. Przygotowanie systemu do pracy

#### 2.1.1. Rozpoczęcie pracy w pakiecie Stata

#### 2.1.2. Zapisywanie wyników pracy z wykorzystaniem pliku *log*

### 2.2. Podstawowe operacje na zbiorach danych

#### 2.2.1. Wczytywanie i zapisywanie zbiorów danych

#### 2.2.2. Importowanie zbiorów danych

#### 2.2.3. Transformacja zmiennych i zbiorów danych

#### 2.2.4. Raportowanie na podstawie zbiorów danych

## **3. Wybrane zagadnienia metody reprezentacyjnej**

### 3.1. Losowanie proste

### 3.2. Losowanie warstwowe

### 3.3. Losowanie systematyczne

## **4. Analiza struktury**

### 4.1. Prezentacja tabelaryczna i graficzna struktury zjawisk

### 4.2. Miary położenia

### 4.3. Stosowanie wag w pakiecie Stata

### 4.4. Obserwacje odstające

## **5. Analiza rozkładu**

### 5.1. Graficzna prezentacja wybranych rozkładów

### 5.2. Badanie zgodności rozkładu empirycznego z rozkładem teoretycznym

## **6. Weryfikacja hipotez statystycznych**

- 6.1. Test istotności dla średniej
- 6.2. Test istotności dla różnicy średnich
- 6.3. Test istotności dla wariancji
- 6.4. Test istotności dla różnicy wariancji
- 6.5. Test istotności dla frakcji
- 6.6. Test istotności dla mediany
- 6.7. Test serii

## **7. Analiza wariancji i kowariancji**

- 7.1. Analiza wariancji z pojedynczą klasyfikacją
- 7.2. Analiza wariancji z podwójną klasyfikacją
- 7.3. Analiza kowariancji

## **8. Analiza korelacji i regresji**

- 8.1. Analiza korelacji
- 8.2. Klasyczny model regresji liniowej
- 8.3. Regresja wieloraka

## **9. Analiza szeregów czasowych**

- 9.1. Przygotowanie zbioru danych do analizy
- 9.2. Graficzna prezentacja szeregu czasowego
- 9.3. Dekompozycja szeregów czasowych
  - 9.3.1. Dekompozycja szeregu bez okresowości
  - 9.3.2. Dekompozycja szeregu addytywnego
  - 9.3.3. Dekompozycja szeregu multiplikatywnego
- 9.4. Modelowanie i prognozowanie szeregu czasowego

## **Bibliografia**

**Załącznik 1. Podstawy programowania w ramach systemu Stata**

**Załącznik 2. Słownik komend - Stata**

**Załącznik 3. Sylabus do przedmiotu Statystyka - program obowiązujący w roku akademickim 2012/2013**