

**ИКОНОМИЧЕСКИ УНИВЕРСИТЕТ - ВАРНА**  
**ФАКУЛТЕТ „ИНФОРМАТИКА“**  
**КАТЕДРА „СТАТИСТИКА И ПРИЛОЖНА МАТЕМАТИКА“**

---

---

Приета от ФС (протокол № 27/ 26.04.2022 г.)

Приета от КС (протокол № 8/ 24.03.2022 г.)

**УТВЪРЖДАВАМ:**

Декан:

(проф. д-р Владимир Сълов)

**У Ч Е Б Н А П Р О Г Р А М А**

ПО ДИСЦИПЛИНАТА: „СТАТИСТИКА ЗА DATA SCIENCE“;

ЗА СПЕЦ: „Data Science“; ОКС „бакалавър“ - редовно обучение

КУРС НА ОБУЧЕНИЕ: 2; СЕМЕСТЪР: 3;

ОБЩА СТУДЕНТСКА ЗАЕТОСТ: 240 ч.; в т.ч. аудиторна 60 ч.

КРЕДИТИ: 8

**РАЗПРЕДЕЛЕНИЕ НА СТУДЕНТСКАТА ЗАЕТОСТ СЪГЛАСНО УЧЕБНИЯ ПЛАН**

<i>ВИД УЧЕБНИ ЗАНЯТИЯ</i>	<i>ОБЩО(часове)</i>	<i>СЕДМИЧНА НАТОВАРЕНОСТ (часове)</i>
АУДИТОРНА ЗАЕТОСТ:		
т. ч.		
• ЛЕКЦИИ	30	2
• УПРАЖНЕНИЯ (семинарни занятия/ лабораторни упражнения)	30	2
ИЗВЪНАУДИТОРНА ЗАЕТОСТ	180	-

Изготвили програмата:

1. ....  
(проф. д-р Веселин Хаджиев)
2. ....  
(гл. ас. д-р Ваня Стоянова)
3. ....  
(гл. ас. д-р Деян Михайлов)

Ръководител катедра: .....  
„Статистика и приложна математика“ (проф. д-р Росен Николаев)

## I. АНОТАЦИЯ

Дисциплината „Статистика за Data Science” е част от фундаменталната подготовка на студентите по специалност Data Science. За усвояване на учебното съдържание е необходимо да се познават основните понятия от линейната алгебра, математическия анализ, теорията на вероятностите и математическата статистика. Целта на обучението по дисциплината е да се развият у студентите следните ключови компетентности в съответствие с препоръката на Съвета на Европейския съюз от 22 май 2018 г.:

- Математическа компетентност и компетентност в областта на точните науки, технологиите и инженерството, изразяващи се в способност и желание за използване на логическо мислене и усвояване на статистическата теория и методология, както и способност за представяне на статистическа информация (формули, модели, концепции, графики и диаграми) и умения за практическо приложение на статистически методи за анализ на масиви от данни.
- Цифрова компетентност, изразяваща се в способност за ползване на цифрова информация и използване на софтуер за приложение на статистическите методи.
- Личностна компетентност, изразяваща се в умение за справяне със сложни задачи и придобиване на способност за организация и постоянство в ученето.

Придобитите знания подпомагат изучаването на останалите специални дисциплини.

Дисциплината е структурирана в петнадесет теми. В тях е включен материал, усвояването на който ще даде възможност на студентите самостоятелно да анализират емпирични данни.

В теми I-IX са включени статистически методи за анализ в различни направления (едномерни и многомерни методи).

Теми X-XIII представят основите на анализа на времевите редове.

В теми XIV-XV са дадени някои от методите за класификация, които намират приложение в различни области и способи, даващи възможност да се намали размерността на данните.

## II. ТЕМАТИЧНО СЪДЪРЖАНИЕ

No. по ред	НАИМЕНОВАНИЕ НА ТЕМИТЕ И ПОДТЕМИТЕ	БРОЙ ЧАСОВЕ		
		Л	СЗ	ЛУ
<b>I. СТАТИСТИЧЕСКО ИЗУЧАВАНЕ</b>		<b>2</b>	<b>2</b>	
1.1.	Обект и предмет на статистическото изучаване			
1.2.	Статистическа информация.			
1.3.	Основни категории в статистиката			
1.4.	Измерителни скали			
<b>II. ПРЕДСТАВЯНЕ НА СТАТИСТИЧЕСКА ИНФОРМАЦИЯ</b>		<b>2</b>	<b>2</b>	
2.1.	Групиране на данни. Статистически редове			
2.2.	Статистически таблици			
2.3.	Графични изображения			
<b>III. ДЕСКРИПТИВНА СТАТИСТИКА</b>		<b>2</b>	<b>2</b>	
3.1.	Емпирични разпределения. Същност и видове			
3.2.	Характеризиране на центъра на разпределението			
3.3.	Измерване на статистическото разсейване			
3.4.	Характеризиране на формата на разпределението			
<b>IV. СТАТИСТИЧЕСКО ОЦЕНЯВАНЕ</b>		<b>2</b>	<b>2</b>	

4.1.	Статистическа извадка - същност и изисквания			
4.2.	Видове оценки.			
4.3.	Интервално оценяване			
<b>V. СТАТИСТИЧЕСКА ПРОВЕВРКА НА ХИПОТЕЗИ</b>		<b>2</b>	<b>2</b>	
5.1.	Параметрични тестове			
5.2.	Непараметрични тестове			
<b>VI. ДИСПЕРСИОНЕН АНАЛИЗ</b>		<b>2</b>	<b>2</b>	
6.1.	Еднофакторен дисперсионен анализ.			
6.3.	Многофакторен дисперсионен анализ.			
<b>VII. КОРЕЛАЦИОНЕН АНАЛИЗ</b>		<b>2</b>	<b>2</b>	
7.1.	Корелационна зависимост.			
7.2.	Параметрични методи.			
7.3.	Непараметрични методи.			
7.4.	Статистическа значимост на корелационните коефициенти			
<b>VIII. ЕДИНИЧНА ЛИНЕЙНА РЕГРЕСИЯ</b>		<b>2</b>	<b>2</b>	
8.1.	Предпоставки за използване и математически модел			
8.2.	Статистическа значимост на модела.			
<b>IX. МНОЖЕСТВЕНА ЛИНЕЙНА РЕГРЕСИЯ</b>		<b>2</b>	<b>2</b>	
9.1.	Предпоставки за използване и математически модел			
9.2.	Анализ на параметрите на модела. Мултиколинеарност.			
<b>X. НЕЛИНЕЙНА РЕГРЕСИЯ</b>		<b>2</b>	<b>2</b>	
10.1.	Нелинейни модели, приводими към линейни.			
10.2.	Нелинеен МНМК.			
10.3.	Логистичен модел.			
<b>XI. ДИНАМИЧНИ (ВРЕМЕВИ) РЕДОВЕ</b>		<b>2</b>	<b>2</b>	
11.1.	Определение за времеви ред. Условия за построяване.			
11.2.	Стационарни редове – характеристики.			
11.3.	Нестационарни редове – характеристики.			
<b>XII. ДЕТЕРМИНИРАНИ КОМПОНЕНТИ НА ДИНАМИЧНИТЕ РЕДОВЕ</b>		<b>2</b>	<b>2</b>	
12.1.	Идентификация на детерминирани компоненти.			
12.2.	Класически методи за изглаждане на динамичните редове.			
12.3.	Адаптивни методи за изглаждане на динамичните редове			
<b>XIII. МОДЕЛИ НА ДИНАМИЧНИТЕ РЕДОВЕ</b>		<b>2</b>	<b>2</b>	
13.1.	Модели на стационарни редове (AR, MA, ARMA).			
13.2.	Модели на нестационарни редове (ARIMA)			
<b>XIV. КЛЪСТЕРЕН АНАЛИЗ</b>		<b>2</b>	<b>2</b>	
14.1.	Същност и задачи на клъстерния анализ.			
14.2.	Функции на разстояние. Разстояние на Минковски и частни случаи. Разстояние на Махаланобис.			
14.3.	Алгоритми за клъстеризация. Метод на k-средни. Йерархичен клъстерен анализ.			
<b>XV. АНАЛИЗ НА ГЛАВНИТЕ КОМПОНЕНТИ</b>		<b>2</b>	<b>2</b>	
15.1.	Същност на метода. Изчисляване и статистическа значимост на главните компоненти			
15.2.	Понижаване на размерността на модела.			
15.3.	Факторен анализ			
<b>Общо:</b>		<b>30</b>	<b>30</b>	

### **III. ФОРМИ НА КОНТРОЛ:**

<b>№. по ред</b>	<b>ВИД И ФОРМА НА КОНТРОЛА</b>	<b>Брой</b>	<b>ИАЗ ч.</b>
<b>1.</b>	<b>Семестриално оценяване</b>		
1.1.	Самостоятелна работа.	1	60
1.2.	Контролна работа	2	40
<b>Общо за семестриалното оценяване:</b>		<b>3</b>	<b>100</b>
<b>2.</b>	<b>Сесийно оценяване</b>		
2.1.	Изпит – тест и решаване на практическа задача	1	80
<b>Общо за сесийното оценяване:</b>		<b>1</b>	<b>80</b>
<b>Общо за всички форми на контрол:</b>		<b>4</b>	<b>180</b>

### **IV. ЛИТЕРАТУРА**

#### **ЗАДЪЛЖИТЕЛНА (ОСНОВНА) ЛИТЕРАТУРА:**

1. Дигитални учебни ресурси по дисциплината, качени в платформата eLearn.
2. Lee, N., Peters, M. Business Statistics. SAGE Publ. Los Angeles, 2016
3. Agresti, A. An Introduction to Categorical Data Analysis. Wiley, Hoboken NJ, 2019
4. James, G., Witten, D. An Introduction to Statistical Learning: with Applications in R. Springer NY, 2021

#### **ПРЕПОРЪЧИТЕЛНА (ДОПЪЛНИТЕЛНА) ЛИТЕРАТУРА:**

1. Mehmetoglu, M., Jacobsen, T. Applied STATISTICS using STATA: A Guide for the Social Scinces SAGE Publications Ltd, 2022
2. Maclnnes, J. An Introduction to Secondary Data Analysis with IBM SPSS Statistics SAGE Publications Ltd, 2017
3. Bruce, P, Bruce, A. Practical Statistics for Data Scientists: 50 Essential Concepts O'Reilly Media, 2017
4. Field, A. Discovering Statistics Using IBM SPSS Statistics SAGE Publications Ltd, 2022
5. Хаджиев, В. Метаанализ и неговата приложимост в икономическите изследвания, Изд. „Наука и икономика“ – ИУ-Варна, Сп. „Известия“ бр.2, 2022
6. Ламбова, М. Статистическата значимост – панацея или препъникамък? Изд. „АзБуки“, Сп. „Математика и информатика“ бр.2, 2021