

Белорусский государственный университет



Многомерный анализ данных

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности:**

1-23 81 08 Медиакоммуникации

Специализация Медиаисследования и социальная аналитика

2019 г.

Учебная программа составлена на основе:
ОСВО 1-23 81 08-2018 и учебного плана Е 23-316/уч. от 11.05.2018

СОСТАВИТЕЛЬ:

О. В. ТЕРЕЩЕНКО, профессор кафедры социальной коммуникации
Белорусского государственного университета, кандидат социологических
наук, доцент

РЕЦЕНЗЕНТ:

Ю. Г. Черняк, заместитель директора Института социологии НАН Беларуси,
канд. социол. наук, доцент

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой социальной коммуникации
(протокол № 13 от 03.05.2019 г.);

Научно-методическим Советом БГУ
(протокол № 5 от 28.06.2019)

Заведующий кафедрой
социальной коммуникации



И.И.Калачёва

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цель учебной дисциплины – формирование у магистрантов теоретических знаний и практических навыков в сфере использования методов многомерного статистического анализа данных в медиаисследованиях и социальной аналитике.

Задачи учебной дисциплины:

1. Познакомить студентов с методами многомерного статистического анализа.
2. Рассмотреть виды статистического моделирования медийных и социальных процессов с использованием статистического программного обеспечения.
3. Применить многомерные статистические методы и модели при решении аналитических задач в сфере медиаисследований и социальной аналитики.

Место учебной дисциплины в системе подготовки специалиста с высшим образованием (магистра).

Учебная дисциплина «Многомерный анализ данных» относится к модулю «Статистический анализ данных в медиаисследованиях» компонент учреждения образования

Учебная дисциплина читается в третьем семестре и имеет непосредственную тематическую и предметную связь с такими учебными дисциплинами как «Анализ данных в социальных науках», «Основы выборочного метода», «Маркетинг в социальных медиа (SMM)».

Требования к компетенциям

Освоение учебной дисциплины «Многомерный анализ данных» должно обеспечить формирование следующих углубленных профессиональных и специальных компетенций.

УК-4 Владеть методологией научного познания и быть готовым использовать её в практической деятельности.

2СК-7 Владеть методами современного анализа данных для решения задач классификации, кластеризации, прогноза.

В результате освоения учебной дисциплины магистрант должен:

знать:

- основные статистические методы обработки и анализа данных;
- методы многомерного моделирования социальных явлений и процессов;
- программное обеспечение статистического моделирования социальных явлений и процессов;

уметь:

- применять на практике алгоритмы многомерного анализа данных медийных и социальных исследований;
- создавать многомерные модели медийных и социальных явлений и процессов;

- интерпретировать полученные в ходе многомерного моделирования результаты;
- владеть:**
- методиками сбора и обработки данных медийных и социальных исследований;
 - алгоритмами многомерного анализа данных;
 - приемами представления результатов исследований.

Структура учебной дисциплины

Дисциплина изучается в 3 семестре. Всего на изучение учебной дисциплины «Многомерный анализ данных» отведено для очной формы получения высшего образования – 108 часов, в том числе 36 аудиторных часов, из них: лекции – 18 часов, семинарские занятия – 14 часов, управляемая самостоятельная работа – 4 часа.

Трудоемкость учебной дисциплины составляет 3 зачетные единицы.

Форма текущей аттестации по учебной дисциплине –зачет.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Тема 1. Корреляционный анализ данных.

Понятие статистической связи между переменными. Направленные и ненаправленные связи. Коэффициент корреляции: интерпретация значений. Меры связи для номинальных переменных. Меры связи для порядковых переменных. Линейная корреляционная связь. Матрица корреляций.

Тема 2. Регрессионный анализ данных.

Зависимые и независимые переменные. Проблема интерколлинеарности. Парная линейная регрессия. Стандартизированные и нестандартизированные коэффициенты регрессии. Множественная линейная регрессия. Логарифмическая и экспоненциальная модели регрессии. Логистическая регрессия. Путевой анализ.

Тема 3. Снижение размерности пространства переменных.

Многомерное пространство переменных. Измерение латентных переменных: семантический дифференциал. Метод главных компонент. Факторный анализ. Многомерное шкалирование. Анализ соответствий. Коэффициенты корреляции

Тема 4. Построение классификаций и типологий.

Понятия классификации и типологизации в социальных исследованиях. Матрица расстояний между объектами в пространстве переменных. Кластерный анализ. Классификация с обучением: линейный дискриминантный анализ.

Тема 5. Когортный анализ.

Понятие и виды статистических когорт. Когорта и поколение. Продольный и поперечный когортный анализ. Факторы возраста, когорты, времени. Смещение эффектов трех факторов. Описательные и экспериментальные когортные исследования.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Дневная (вечерняя) форма получения образования с применением дистанционных образовательных технологий

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСП	Форма контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
	Корреляционный анализ данных Понятие статистической связи между переменными. Направленные и ненаправленные связи. Коэффициент корреляции: интерпретация значений. Меры связи для номинальных переменных. Меры связи для порядковых переменных. Линейная корреляционная связь. Матрица корреляций	4		2 2 (ДО)				Эссе и устная защита Задание на образовательном портале LMS MOODLE
2	Регрессионный анализ данных Зависимые и независимые переменные. Парная линейная регрессия. Стандартизированные и нестандартизированные коэффициенты регрессии. Множественная линейная регрессия. Логарифмическая и экспоненциальная модели регрессии. Логистическая регрессия. Путевой анализ.	4		2				Решение задач и обсуждение результатов: доклад и устная защита

3	Снижение размерности пространства переменных Многомерное пространство переменных. Измерение латентных переменных: семантический дифференциал. Метод главных компонент. Факторный анализ. Многомерное шкалирование. Анализ соответствий. Коэффициенты корреляции	4		2			2	Работа с кейсами
4	Построение классификаций и типологий Понятия классификации и типологизации в социальных исследованиях. Матрица расстояний между объектами в пространстве переменных. Кластерный анализ. Классификация с обучением: линейный дискриминантный анализ.	4		2			2	Доклад с презентацией
5	Когортный анализ Понятие и виды когорт. Когорта и поколение. Продольный и поперечный когортный анализ. Факторы возраста, когорты, времени. Смещение эффектов трех факторов. Описательные и экспериментальные когортные исследования.	2		2 2 (ДО)				Отчет с последующей устной защитой Тест на образовательном портале LMS MOODLE
	ИТОГО	18		14			4	

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Волкова, П. А. Статистическая обработка данных в учебно-исследовательских работах: Учебное пособие / П. А. Волкова, А. Б. Шипунов. – М.: Форум, 2016. – 96 с.
2. Дадян, Э. Г. Методы, модели, средства хранения и обработка данных: Учебник / Э. Г. Дадян, Ю. А. Зеленков. – М.: Вузовский учебник, ИНФРА-М, 2017. – 168 с.
3. Максимов, С. И. Статистический анализ и обработка данных с применением MS Excel и SPSS: учебно-методическое пособие / С. И. Максимов, Е. М. Зайцев, Е. И. Князева – Минск. РИВШ, 2012. – 112с.
4. Миркин, Б. Г. Введение в анализ данных: учебник и практикум для бакалавриата и магистратуры / Б. Г. Миркин. – М.: Юрайт, 2017. – 176 с.
5. Терещенко, О. В. Многомерный статистический анализ данных в социальных науках: учеб. Пособие (допущено Министерством образования) / О. В. Терещенко, Н. В. Курилович, Е. И. Князева. – Минск: БГУ, 2012. – 239 с.

Перечень дополнительной литературы

1. Айвазян С. А. Прикладная статистика в задачах и упражнениях: Учебник для вузов / С. А. Айвазян, В. С. Мхитарян. – М.: ЮНИТИ-ДАНА, 2001. – 270 с.
2. Афифи, А. Статистический анализ: подход с использованием ЭВМ. / А. Афифи, С. Эйзен. – М.: Мир, 1982. – 488 с.
3. Бессокирная, Г. П. Дискриминантный анализ для отбора информативных переменных // Социология: методология, методы и математическое моделирование (Социология: 4М). – 2003, № 16. – С. 25–35.
4. Боровиков В. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. / В. Боровиков. – СПб.: Питер, 2003. – 688с.
5. Елисеева, И. И. Фиктивные переменные в анализе данных / И. И. Елисеева, С. В. Курышева // Социология: методология, методы и математическое моделирование (Социология: 4М). – 2010, № 30. – С. 43–63.
6. Елисеева, И. И. Основные процедуры многомерного статистического анализа / И. И. Елисеева, Е. В. Семенова. – Л.: УЭФ, 1993. – 78 с.
7. Ким, О. Дж. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка. – М.: Финансы и статистика, 1989. – 215 с.
8. _____, А. О. Анализ социологических данных с помощью пакета SPSS: учеб. пособие для _____ . – М. : ГУ ВШЭ, 2006. – 281 с.
9. Наследов, А. IBM SPSS 20 и AMOS: Профессиональный статистический анализ данных / А. Наследов. – СПб: Питер, 2013. – 416 с.

10. Татарова, Г. Г. Основы типологического анализа в социологическом исследовании / Г. Г. Татарова. – М.: Высшее образование и наука, 2007. – 236 с.

11. Терещенко, О. В., Королева, И. Исследования социального неравенства: процессуальный подход / О. В. Терещенко, И. Королева // Социология и общество: социальное неравенство и социальная справедливость (Екатеринбург, 19–21 октября 2016 года) [Электронный ресурс] Материалы V Всероссийского социологического конгресса / отв. ред. В. А. Мансуров – Электрон. диск. – М.: Российское общество социологов, 2016. – С. 10427–10436

12. Терещенко, О. В. Метод когортного анализа в социальных исследованиях / Социология: методология, методы и математическое моделирование (Социология: 4М). – 2009, № 29. – С. 172–185.

13. Терещенко, О. В. Прикладная статистика для социальных наук: компьютерный практикум для студентов гуманитарных специальностей. / О. В. Терещенко. – Минск, БГУ, 2002. – 96 с.

Перечень рекомендуемых средств диагностики и методика формирования итоговой оценки

Для диагностики компетенций могут использоваться следующие формы:

- устная – доклады на семинарских занятиях, устный зачет;
- письменная – работа с кейсами;
- устно-письменная – отчеты по домашним практическим заданиям с их устной защитой, проект;
- техническая – электронный тест, презентация.

Оценка за ответы на семинарских занятиях может включать в себя полноту ответа, наличие аргументов, примеров из практики и т.д.

При оценивании доклада обращается внимание на: содержание и полноту раскрытия темы, структуру и последовательность изложения, источники и их интерпретацию, корректность оформления и т.д.

При оценивании отчета по домашним практическим заданиям (контент-анализ) с их устной защитой обращается внимание на: структуру и содержание программы исследования, корректность формирования выборочной совокупности, наличие кодировочной таблицы, наличие файла с введенными данными (SPSS, Excel, др.), полноту аналитической записки.

При оценивании работы с кейсами обращается внимание на активность работы всех членов группы, быстроту выполнения заданий, краткость и четкость изложения, этику ведения дискуссии, отбор информации.

При оценивании презентации обращается внимание на: формальные требования к построению компьютерной презентации, содержательность выступления, продолжительность выступления.

Оценка проекта может включать актуальность исследуемой проблемы, корректность используемых методов исследования, привлечение знаний из различных областей, организация работы группы, практикоориентированность полученных результатов.

Оценивание электронного теста осуществляется следующим образом: общая сумма баллов за все правильные ответы составляет наивысший балл, например, 100 баллов. Устанавливается диапазон баллов, которые необходимо набрать для того, чтобы получить отличную («9», «10»), хорошую («6», «7», «8»), удовлетворительную («5», «4») или неудовлетворительную оценки («3», «2», «1»).

Формой текущей аттестации по дисциплине «Многомерный анализ данных» учебным планом предусмотрен зачет.

При формировании итоговой оценки используется рейтинговая оценка знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине.

Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний и текущей аттестации в рейтинговую оценку:

Формирование оценки за текущую успеваемость:

- ответы на семинарских занятиях – 25 %;
- выполнение заданий ДО – 25 %;
- решение задач и обсуждение результатов – 25 %.
- Тест на образовательном портале LMS MOODLE – 25 %

Примерный перечень заданий для управляемой самостоятельной работы студентов

Тема 3. Снижение размерности пространства переменных

Выбрать три рекламных материала (кейсы). Оценить каждый из них по шкалам семантического дифференциала.

Вычислить значения шкал семантического пространства и выполнить факторный анализ с использованием программного средства. Измерение латентных переменных: семантический дифференциал. Проинтерпретировать сходства и различия полученных семантического и факторного пространств.

(Форма контроля – работа с кейсами).

Тема 4. Построение классификаций и типологий

Выбрать один из предложенных наборов данных. Сформулировать гипотезу о структуре представленной в наборе данных выборки. Проверить гипотезу методом дискриминантного анализа.

Проанализировать структуру той же выборки методом кластерного анализа. Сравнить и проинтерпретировать результаты, полученные двумя методами.

(Форма контроля – доклад с презентацией)

Примерная тематика семинарских занятий

Семинар № 1. Виды коэффициентов корреляции в зависимости от измерения переменных

(Форма контроля – эссе и устная защита).

Семинар № 2. Применение регрессионных моделей в исследованиях социальной мобильности и социального неравенства.

(Форма контроля – доклад и его защита).

Семинар № 3. Анализ структуры матрицы корреляций между переменными.

(Форма контроля – работа с кейсами).

Семинар №4. Построение классификации методом группировки

(Форма контроля – презентация).

Семинар №5. Когортный анализ проникновения интернета в Беларуси за последние 20 лет (по данным Белстата)

(Форма контроля – отчет с последующей устной защитой).

Семинарские занятия – дистанционная форма

№	Виды семинарских заданий	Формы контроля знаний	Обеспечение на образовательном портале
1	Коллоквиум по теме «Особенности статистических связей»	форум на портале LMS MOODLE	Список тем, литература
2	Итоговый тест на портале LMS MOODLE	Тест на портале LMS MOODLE	Компьютерный тест

Описание инновационных подходов и методов к преподаванию учебной дисциплины (эвристический, проективный, практико-ориентированный, метод анализа конкретных ситуаций)

При организации образовательного процесса используются следующие подходы:

практико-ориентированный подход, который предполагает:

- освоение содержания образования через решения практических задач;
- приобретение навыков эффективного выполнения разных видов профессиональной деятельности;
- ориентацию на генерирование идей, реализацию групповых студенческих проектов, развитие предпринимательской культуры;
- использованию процедур, способов оценивания, фиксирующих сформированность профессиональных компетенций.

метод анализа конкретных ситуаций (кейс-метод), который предполагает:

- приобретение студентом знаний и умений для решения практических задач;
- анализ ситуации, используя профессиональные знания, собственный опыт, дополнительную литературу и иные источники.

метод проектного обучения, который предполагает:

- способ организации учебной деятельности студентов, развивающий актуальные для учебной и профессиональной деятельности навыки планирования, самоорганизации, сотрудничества и предполагающий создание собственного продукта;
- приобретение навыков для решения исследовательских, творческих, социальных, предпринимательских и коммуникационных задач.

Методические рекомендации по организации самостоятельной работы обучающихся, кроме подготовки к экзамену, подготовка к зачету

При изучении учебной дисциплины рекомендуется использовать следующие формы самостоятельной работы: – поиск (подбор) и обзор литературы и электронных источников по индивидуально заданной проблеме курса; – выполнение домашнего задания (проекта); – работы, предусматривающие решение конкретных ситуаций и выполнение

упражнений, выдаваемых на практических занятиях; – изучение материала, вынесенного на самостоятельную проработку; – подготовка к практическим семинарским занятиям; – подготовка и написание докладов, эссе и презентаций на заданные темы.

Примерный перечень тем для докладов

1. Инновации как фактор социального развития.
2. Социальные инновации как процесс управления жизнедеятельностью общества.
3. Социальные инновации XXI века.
4. Инновации в культуре.
5. Поддержка социальных инноваций в ЕС.
6. Роль политических инноваций в условиях обновления общества.
7. Технологические инновации как потребность социально-экономического развития общества.
8. Инновационная экономика.
9. Экологические инновации.
10. Организационно-управленческие инновации.
11. Маркетинговые инновации как направление инновационной политики.
12. Информационные инновации.

Примерный перечень вопросов к зачету

1. Понятие и особенности статистической связи
2. Меры связи для номинальных переменных
3. Меры связи для порядковых переменных
4. Меры связи для количественных мер связи
5. Корреляционные и причинные связи
6. Матрица корреляций
7. Зависимые и независимые переменные
8. Проблема интерколлинеарности
9. Парная линейная регрессия
10. Множественная линейная регрессия
11. Логарифмическая, экспоненциальная, логистическая модели регрессии
12. Путевой анализ
13. Многомерное пространство переменных
14. Метод главных компонент
15. Факторный анализ
16. Метод семантического дифференциала
17. Типология и классификация социальных объектов.
18. Матрица расстояний между объектами.
19. Кластерный анализ.
20. Линейный дискриминантный анализ
21. Понятие и виды когорт

22. Продольный и поперечный когортный анализ
23. Возраст и время в когортных исследованиях
24. Описательные и экспериментальные когортные исследования

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Компьютерный анализ данных	Кафедра социальной коммуникации	нет	13 от 03.05.2019 г
2. Язык программирования R в социальных науках	Кафедра социальной коммуникации	нет	13 от 03.05.2019 г
3. Инфографика	Кафедра социальной коммуникации	нет	13 от 03.05.2019 г

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО
ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 201_ г.)

Заведующий кафедрой

УТВЕРЖДАЮ
Декан факультета
