

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION

**Federal State Autonomous Educational Institution of Higher Education
«National Research Lobachevsky State University of Nizhny Novgorod»**

Факультет социальных наук

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Working programme of the discipline

Basics of Big Data and Data Science

Higher education level

Master degree

Area of study / speciality

37.04.01 - Psychology

Focus /specialization of the study programme

Cyberpsychology

Mode of study

full-time

Nizhny Novgorod

Year of commencement of studies 2024

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.16 Основы Big Data и Data Science относится к части, формируемой участниками образовательных отношений образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства	
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	Для текущего контроля успеваемости	Для промежуточной аттестации
ПК-1.2: Способен в ходе психологического исследования собирать данные, их статистически обрабатывать и анализировать	<p>ПК-1.2.1: Знает методы сбора данных в психологии и математической статистики.</p> <p>ПК-1.2.2: Умеет обрабатывать данные с помощью математической статистики, анализировать на их основе результаты психологического исследования.</p> <p>ПК-1.2.3: Владеет методами и средствами математической статистики и анализа данных.</p>	<p>ПК-1.2.1: Знать: 1. Методы сбора больших данных. 2. Методы сбора данных для задачи их анализа и интерпретации.</p> <p>ПК-1.2.2: Уметь: 1. Обрабатывать большие данные. 2. Интерпретировать большие данные.</p> <p>ПК-1.2.3: Владеть: 1. Методами статистики для интерпретации данных. 2. Средствами статистики для интерпретации данных.</p>	Кейс-задание	<p>Зачёт: Тест Задания Контрольные вопросы</p>
ДПК-4.3: Способен использовать современные информационные технологии в психологической работе	<p>ДПК-4.3.1: Знает назначение и принципы работы ключевых информационных технологий и компьютерных программ, используемых в психологических исследованиях.</p> <p>ДПК-4.3.2: Умеет применять в практике психологических исследований современные информационные</p>	<p>ДПК-4.3.1: Знать: 1. Назначение программ для работы с данными и большими данными. 2. Принципы работы программ для работы с данными и большими данными.</p> <p>ДПК-4.3.2: Уметь: 1. Применять на практике</p>	Кейс-задание	<p>Зачёт: Тест Задания Контрольные вопросы</p>

	технологии. ДПК-4.3.3: Владеет ключевыми компьютерными программами, используемыми в психологических исследованиях.	технологии сбора больших данных. 2. Применять на практике технологии анализа больших данных. ДПК-4.3.3: Владеть: 1. Программами сбора больших данных. 2. Программами анализа больших данных.		
--	--	--	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	3
Часов по учебному плану	108
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	0
- занятия семинарского типа (практические занятия / лабораторные работы)	36
- КСР	1
самостоятельная работа	71
Промежуточная аттестация	0 Зачёт

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

Наименование разделов и тем дисциплины	Всего (часы)	в том числе			Самостоятельная работа обучающегося, часы
		Контактная работа (работа во взаимодействии с преподавателем), часы из них			
		Занятия лекционного типа	Занятия семинарского типа (практические занятия/ лабора торные работы), часы	Всего	
	0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0
Тема 1. Python для работы с данными.	24		8	8	16

Тема 2. SQL для работы с Big Data.	25		8	8	17
Тема 3. Git и AirFlow.	29		10	10	19
Тема 4. Big Data в исследовании социальных данных.	29		10	10	19
Аттестация	0				
КСР	1			1	
Итого	108	0	36	37	71

Contents of sections and topics of the discipline

Тема 1. Python для работы с данными.

Тема 2. SQL для работы с Big Data.

Тема 3. Git и AirFlow.

Тема 4. Big Data в исследовании социальных данных.

Тема 1. Python for Data Processing.

Тема 2. SQL for Big Data Management.

Тема 3. Git and AirFlow.

Тема 4. Big Data in Social Data Research.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведенным в п. 5.

По дисциплине предусмотрено выполнение самостоятельной работы в виде анализа конкретного случая и попытки решения проблемы, описанной в кейсе.

По теме кейса студентом самостоятельно собирается материал, обрабатывается, оформляется в виде письменной работы и презентации. Работы обсуждаются на практических занятиях. В ходе группового обсуждения рассматривается та или иная проблема виртуальной реальности. На всех этапах выполнения работы преподаватель оказывает консультативную помощь и методическое обеспечение.

5. Assessment tools for ongoing monitoring of learning progress and interim certification in the discipline (module)

5.1 Model assignments required for assessment of learning outcomes during the ongoing monitoring of learning progress with the criteria for their assessment:

5.1.1 Model assignments (assessment tool - Case task) to assess the development of the competency ПК-1.2:

1. Проблема больших данных в гуманитарных науках.
2. Применение больших данных при решении прикладных задач.
3. Роль больших данных в современном мире.

5.1.2 Model assignments (assessment tool - Case task) to assess the development of the competency ДПК-4.3:

1. Анализ социальных медиа-данных для идентификации и понимания психологических трендов и эмоциональных реакций пользователей.
2. Разработка модели машинного обучения для прогнозирования психологических состояний, таких как депрессия или тревога, на основе собранных данных.
3. Исследование влияния факторов окружающей среды, таких как погода или временные изменения, на психологическое состояние людей с помощью анализа больших данных.

Assessment criteria (assessment tool — Case task)

Grade	Assessment criteria
pass	Минимально допустимый уровень знаний. Допущено много негрубых ошибки. Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме. Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами
fail	Уровень знаний ниже минимальных требований. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.

5.2. Description of scales for assessing learning outcomes in the discipline during interim certification

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатора достижения компетенций)	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено			зачтено			
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено много негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений	При решении стандартных задач не продемонстрированы основные	Продемонстрированы основные умения. Решены типовые	Продемонстрированы все основные умения. Решены все	Продемонстрированы все основные умения. Решены все	Продемонстрированы все основные умения. Решены	Продемонстрированы все основные умения. Решены все основные

	вследствие отказа обучающегося от ответа	умения. Имели место грубые ошибки	задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме	основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами	все основные задачи с отдельным и несущественными недочетами, выполнены все задания в полном объеме	задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие базовых навыков. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач с некоторым и недочетами	Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов	Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов	Продемонстрирован творческий подход к решению нестандартных задач

Scale of assessment for interim certification

Grade		Assessment criteria
pass	outstanding	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "outstanding", the knowledge and skills for the relevant competencies have been demonstrated at a level higher than the one set out in the programme.
	excellent	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "excellent",
	very good	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "very good",
	good	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "good",
	satisfactory	All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "satisfactory", with at least one competency developed at the "satisfactory" level.
fail	unsatisfactory	At least one competency has been developed at the "unsatisfactory" level.
	poor	At least one competency has been developed at the "poor" level.

5.3 Model control assignments or other materials required to assess learning outcomes during the interim certification with the criteria for their assessment:

5.3.1 Model assignments (assessment tool - Test) to assess the development of the competency ПК-1.2

Инструкция: Выберите один или несколько вариантов ответа.

1. Для работы с большими данными используются/используется:

1. поправки на множественность сравнений
2. бутстрап
3. непараметрическая статистика
4. параметрическая статистика

2. Для визуализации Big Data можно использовать:

1. Tableau
2. Excel
3. SPSS
4. Padlet

3. Какой метод используется для обработки и анализа Big Data в психологии?

1. Машинное обучение
2. Статистический анализ
3. Кластерный анализ
4. Нейронные сети

5.3.2 Model assignments (assessment tool - Test) to assess the development of the competency ДПК-4.3

Инструкция: Выберите один вариант ответа.

1. Какая методика может использоваться для предобработки данных в Data Science?

1. Удаление выбросов
2. Заполнение пропущенных значений
3. Масштабирование данных
4. Все вышеперечисленное

2. Какая задача Data Science связана с группировкой объектов на основе их схожести?

1. Классификация
2. Кластеризация
3. Регрессия
4. Ассоциативный анализ

3. Какой тип алгоритма машинного обучения используется для предсказания категории или класса новых данных?

1. Решающие деревья
2. Нейронные сети
3. К-ближайших соседей
4. Логистическая регрессия

Assessment criteria (assessment tool — Test)

Grade	Assessment criteria
pass	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция

Grade	Assessment criteria
	сформирована на уровне «удовлетворительно»
fail	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

5.3.3 Model assignments (assessment tool - Assignments) to assess the development of the competency ПК-1.2

Задание 1

Анализировать по материалам кейса самостоятельной работы статистические закономерности

Задание 2

Описать методы статистики по материалам кейса самостоятельной работы

Задание 3

Описать правила организации сбора больших данных

5.3.4 Model assignments (assessment tool - Assignments) to assess the development of the competency ДПК-4.3

Задание 1

Проанализируйте влияние социальных медиа на психологические процессы с использованием Big Data.

Задание 2

Исследуйте роль Data Science в анализе ментального здоровья и диагностике психических расстройств.

Задание 3

Соберите и проанализируйте данные о потребительском поведении для определенной отрасли.

Assessment criteria (assessment tool — Assignments)

Grade	Assessment criteria
pass	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
fail	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

5.3.5 Model assignments (assessment tool - Control questions) to assess the development of the competency ПК-1.2

1. Методы сбора больших данных.
2. Методы сбора данных для задачи их анализа и интерпретации.
3. Роль больших данных в современном мире.
4. Роль больших данных в гуманитарных исследованиях.
5. Типы больших данных.
6. Понятие Data Science.
7. Понятие Data Analysts.
8. Современные тренды в количественной оценке социальных феноменов.
9. Применение SQL для анализа больших данных.
10. Применение AirFlow в Data Science.
11. Применение Python в Data Science.
12. Применение Git в Data Science.
13. Средства визуализации больших данных.
14. Статистические методы работы большими данными.
15. Способы и технологии хранения больших данных.
16. Применение методов анализа больших данных в киберпсихологии.
17. Что такое Big Data и как оно применяется в психологии?
18. Какие основные источники данных используются в психологических исследованиях?
19. Какие методы обработки и анализа Big Data применяются в психологии?
20. Какие преимущества и ограничения существуют при использовании Big Data в психологических исследованиях?
21. Какие навыки и знания необходимы для работы с Big Data в психологии?
22. Какие этические вопросы возникают при использовании Big Data в психологии и как они решаются?
23. Какие инструменты и программное обеспечение используются для обработки и анализа Big Data в психологии?
24. Какие типы анализа данных можно провести с помощью Big Data в психологии?
25. Какие вызовы и проблемы возникают при сборе и хранении больших объемов данных в психологических исследованиях?
26. Какие принципы и методы принятия решений применяются при анализе Big Data в психологии?
27. Какие тенденции и перспективы развития Big Data в психологии?
28. Какие примеры успешного применения Big Data в психологии можно назвать?
29. Какие методы визуализации данных используются при анализе Big Data в психологии?
30. Какие метрики и показатели используются для оценки эффективности анализа Big Data в психологии?

5.3.6 Model assignments (assessment tool - Control questions) to assess the development of the competency ДПК-4.3

1. Какие типы машинного обучения применяются в анализе Big Data в психологии?
2. Какие методы фильтрации и предварительной обработки данных используются при работе с Big Data в психологии?
3. Какие проблемы и сложности возникают при интерпретации результатов анализа Big Data в психологии?
4. Какие роли и профессии связаны с работой с Big Data в психологии?
5. Какие методы группировки и классификации данных применяются в анализе Big Data в психологии?
6. Какие модели прогнозирования используются при работе с Big Data в психологии?
7. Какие примеры задач и исследований могут быть решены с помощью Big Data в психологии?
8. Какие методы извлечения знаний и паттернов используются при анализе Big Data в психологии?
9. Какие статистические методы и тесты применяются для проверки гипотез при работе с Big Data в психологии?
10. Какие методы сегментации и разделения данных используются в анализе Big Data в психологии?
11. Какие методы восстановления и очистки данных применяются при работе с Big Data в психологии?
12. Какие примеры исследований социальных сетей и анализа текстовых данных можно назвать в контексте Big Data в психологии?
13. Какие техники и методы визуализации данных используются при работе с Big Data в психологии?
14. Какие принципы и методы анонимизации данных используются для защиты конфиденциальности в анализе Big Data в психологии?
15. Какие примеры исследований использования Big Data для предсказания поведения и предпочтений людей в психологии?
16. Какие примеры исследований в области медицинской психологии и биометрики используют Big Data?
17. Какие методы кластерного анализа и ассоциативных правил применяются в анализе Big Data в психологии?
18. Какие методы фильтрации и поиска информации используются при работе с Big Data в психологии?
19. Какие примеры задачи обнаружения аномалий и выбросов могут быть решены с помощью Big Data в психологии?
20. Какие принципы и методы обеспечения качества данных используются в анализе Big Data в психологии?
21. Какие методы оптимизации и ускорения анализа Big Data применяются в психологии?
22. Какие примеры исследований использования Big Data для прогнозирования психологических и эмоциональных состояний людей?
23. Какие методы временного анализа и прогнозирования используются при работе с Big Data в психологии?
24. Какие примеры исследований использования Big Data в области психофизиологии и биометрии?
25. Какие методы объединения и интеграции данных используются при работе с Big Data в психологии?
26. Какие принципы и методы обработки текстовых данных применяются в анализе Big Data в психологии?
27. Какие примеры задачи предсказания и ранжирования событий и результатов могут быть решены с помощью Big Data в психологии?

28. Какие методы регрессионного анализа и моделирования используются при работе с Big Data в психологии?

29. Какие примеры исследований использования Big Data в области психологии потребительского поведения и маркетинга?

30. Какие методы сжатия и хранения данных применяются при работе с Big Data в психологии?

Assessment criteria (assessment tool — Control questions)

Grade	Assessment criteria
pass	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
fail	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Келлехер Джон. Наука о данных: базовый курс : Научно-популярная литература. - Москва : ООО "Альпина Паблишер", 2020. - 222 с. - ВО - Магистратура. - ISBN 978-5-9614-3170-4., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=740846&idb=0>.
2. Анализ данных : учебник / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. - Москва : Юрайт, 2023. - 490 с. - (Высшее образование). - ISBN 978-5-534-00616-2. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847079&idb=0>.
3. Советов Б. Я. Базы данных : учебник / Б. Я. Советов, В. В. Цехановский, В. Д. Чертовской. - 3-е изд. ; пер. и доп. - Москва : Юрайт, 2023. - 420 с. - (Высшее образование). - ISBN 978-5-534-07217-4. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847017&idb=0>.
4. Миркин Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. - Москва : Юрайт, 2023. - 174 с. - (Высшее образование). - ISBN 978-5-9916-5009-0. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=843081&idb=0>.
5. Толстобров А. П. Управление данными : учебное пособие / А. П. Толстобров. - 3-е изд. ; пер. и доп. - Москва : Юрайт, 2023. - 272 с. - (Высшее образование). - ISBN 978-5-534-14162-7. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=848160&idb=0>.

Дополнительная литература:

1. Баюк Д.А. Практическое применение методов кластеризации, классификации и аппроксимации на основе нейронных сетей : монография / Баюк Д.А.; Баюк О.А.; Берзин Д.В. - Москва : Прометей, 2020. - 448 с. - ISBN 978-5-00172-086-7., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=808002&idb=0>.
2. Маркин А. В. Программирование на SQL в 2 ч. Часть 1 : учебник и практикум / А. В. Маркин. -

- 3-е изд. ; пер. и доп. - Москва : Юрайт, 2023. - 429 с. - (Высшее образование). - ISBN 978-5-534-15817-5. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847370&idb=0>.
3. Маркин А. В. Программирование на SQL в 2 ч. Часть 2 : учебник и практикум / А. В. Маркин. - 3-е изд. ; испр. и доп. - Москва : Юрайт, 2023. - 385 с. - (Высшее образование). - ISBN 978-5-534-15818-2. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=845724&idb=0>.
4. Шапцев В. А. Теория информации. Теоретические основы создания информационного общества / Шапцев В. А., Бидуля Ю. В. - Москва : Юрайт, 2022. - 177 с. - (Высшее образование). - URL: <https://urait.ru/bcode/490739> (дата обращения: 05.01.2022). - ISBN 978-5-534-02989-5 : 499.00. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=784293&idb=0>.
5. Зараменских Е. П. Основы бизнес-информатики : учебник и практикум / Е. П. Зараменских. - 2-е изд. - Москва : Юрайт, 2023. - 470 с. - (Высшее образование). - ISBN 978-5-534-15039-1. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=839813&idb=0>.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

Интернет-ресурсы:

1. <http://opendata.university> - Сайт Университетского консорциума исследователей больших данных.
2. <http://рпо.рф/> - Сайт Российского психологического общества. Содержит материалы по многим направлениям психологии.
3. <http://www.ipras.ru> - Сайт Института психологии РАН.
4. <http://www.mtu-net.ru> - Ресурс, созданный сотрудниками Психологического института им. Л.Г. Щукиной.
5. <http://psyberia.ru/> - Образовательный психологический проект. Представлены разнообразные информационные материалы по многим направлениям психологии.
6. <http://psyjournals.ru/> - Крупнейший в Интернете Портал психологических изданий.
7. <http://www.psystudy.com/> - Мультидисциплинарный научный психологический интернет-журнал "Психологические исследования" публикует оригинальные статьи по различным отраслям психологии и смежных наук.

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению подготовки/специальности 37.04.01 - Psychology.

Author(s): Демарева Валерия Алексеевна, кандидат психологических наук.

Заведующий кафедрой: Голубин Роман Викторович, кандидат исторических наук.

Программа одобрена на заседании методической комиссии от 17.12.2023, протокол № 7.