

**Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА  
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ  
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

---

Институт общественных наук

*(наименование института)*

Кафедра теоретической социологии и эпистемологии

*(наименование кафедры)*

УТВЕРЖДЕНА

кафедрой теоретической  
социологии и эпистемологии

Протокол от «10» июня 2020 г.

№5

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Б1.В.ДВ.01.23.03 Инструменты работы с данными

*(индекс и наименование дисциплины)*

Minor "Современные методы анализа данных"

38.03.02 Менеджмент

*(код, наименование направления подготовки)*

Стратегическое управление компанией (Liberal Arts)

*направленность (профиль)*

бакалавр

*(квалификация)*

очная

*(форма обучения)*

Год набора - 2021

Москва, 2020 г.

**Автор–составитель:**

Кандидат социологических наук, доцент кафедры теоретической социологии и эпистемологии, Астахова А.С.

Заведующий кафедрой теоретической социологии и эпистемологии, декан философско-социологического факультета, к.с.н. Вахштайн В.С.

## Содержание

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы	4
2. Объем и место дисциплины в структуре ОП ВО	4
3. Содержание и структура дисциплины	5
4. Материалы текущего контроля и фонд оценочных средств промежуточной аттестации по дисциплине	6
5. Методические указания для обучающихся по освоению дисциплины	9
6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	10
6.1. Основная литература.	10
6.2. Дополнительная литература.	10
6.3. Учебно-методическое обеспечение самостоятельной работы.	10
6.4. Нормативные правовые документы.	10
6.5. Интернет-ресурсы, справочные системы.	11
6.6. Иные рекомендуемые источники.	11
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	11

## 1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина Б1.В.ДВ.01.23.03 «Инструменты работы с данными» обеспечивает овладение следующими компетенциями с учетом индикатора:

Код компетенции	Наименование компетенции	Код индикатора компетенции	Наименование индикатора компетенций
СК ОС LA - 25	Способен осуществлять подготовку, анализ, интерпретацию данных из первичных и вторичных источников, а также визуализировать и представлять результаты исследования	СК ОС LA-25.1.	Представляет результаты проведенного анализа данных, может описать и объяснить происхождение полученных значений; демонстрирует владение специализированными программами для работы с данными;

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

Код индикатора компетенции	Результаты обучения
СК ОС LA-25.1	на уровне знаний: <ul style="list-style-type: none"> <li>• основ математической статистики;</li> <li>• основ анализа данных;</li> </ul> основных принципов сбора, подготовки, хранения данных;
	на уровне умений: <ul style="list-style-type: none"> <li>• осуществлять ввод и «чистку» данных;</li> </ul> отбирать и анализировать вторичные данные;
	на уровне навыков: <ul style="list-style-type: none"> <li>• пользоваться справочными информационными системами для поиска дополнительной информации;</li> </ul> использования специализированных статистических пакетов для анализа данных;

## 2. Объем и место дисциплины в структуре ОП ВО

- Общая трудоемкость дисциплины: 108ак./81астр. часов (3 ЗЕТ);
- 30/22,5 ак./астр. часов выделено на контактную работу с преподавателем и 42ак./31,5 /астрон. часа на самостоятельную работу обучающихся;
- Б1.В.ДВ.01.23.03. Инструменты работы с данными осваивается на 3 курсе обучения (6 семестр).
- дисциплина реализуется после изучения: Б1.О.10 Основы математики, Б1.О.11 Информационные технологии, Б1.В.ДВ.01.23.02 Введение в работу с данными: источники, отбор, подготовка.

### 3. Содержание и структура дисциплины

Таблица 1.

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.						Форма текущего контроля успеваемост и <sup>4</sup> , промежуютч ной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КС Р		
1	Инструменты работы с данными: введение	10/7,5			4/3		6/4,5	О
2	Основные принципы работы с R	12/9			4/3		8/6	О, ПЗ
3	Факторы и таблицы в R	12/9			4/3		8/6	О, ПЗ
4	Графики в R	12/9			4/3		8/6	О, ПЗ
5	Инструменты работы с Big Data	24/18			12/9		12/9	О, ПЗ
Консультация		2/1,5						
Промежуточная аттестация		36/27						Экзамен
Всего:		108/81			28/21		42/31,5	

Примечание: \*\* – формы текущего контроля успеваемости: опрос (О), практические задания (пз)

### Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
Тема 1	Инструменты работы с данными: введение	Классификации данных. Типы инструментов для работы с данными. Основные концепции организации реляционных баз данных и хранилища данных (DWH) Отличие баз данных от хранилищ данных (Digital WareHouses - DWH). Подходы к организации хранилищ данных и инструменты их реализации. Назначение хранилищ данных и их место в корпоративной экосистеме. Сырые (raw) и нормализованные данные.
Тема 2	Основные принципы работы с R	Объектно-ориентированное программирование. Векторы. Матрицы и массивы. Кадры данных. Арифметические и логические операторы и значения.
Тема 3	Факторы и таблицы в R	Факторы и уровни. Типичные функции, используемые с факторами.
Тема 4	Графики в R	Управление графическими окнами. Графические функции. Команды управления графиками нижнего уровня.

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
Тема 5	Инструменты работы с Big Data	Принципы и подходы распределенной обработки больших данных. Особенности HDFS. Модель распределенных вычислений MapReduce. Экосистема Hadoop. Фреймворк Apache Spark и его отличия от других баз данных. Распределенная обработка данных и управление распределенными данными.

#### **4. Материалы текущего контроля и фонд оценочных средств промежуточной аттестации по дисциплине**

##### **4.1 Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации.**

**4.1.1. В ходе реализации дисциплины Б1.В.ДВ.01.23.03. «Инструменты работы с данными» используются следующие методы текущего контроля и успеваемости обучающихся:**

– при проведении практических занятий:

- Опрос
- Практические задания

**4.1.2. Экзамен проводится с применением следующих методов (средств): устный экзамен, состоящий из ответов на 2 вопроса.**

##### **4.2. Материалы текущего контроля успеваемости.**

**Примерные вопросы для подготовки к практическим занятиям по темам:**

##### **Тема 1. Инструменты работы с данными: введение**

1. Классификации данных. Типы инструментов для работы с данными.

##### **Тема 2. Основные принципы работы с R**

1. Отличие баз данных от хранилищ данных (Digital WareHouses - DWH).
2. Сырые (raw) и нормализованные данные.
3. Объектно-ориентированное программирование.

##### **Тема 3. Факторы и таблицы в R**

1. Векторы.
2. Матрицы и массивы.
3. Кадры данных.
4. Арифметические и логические операторы и значения.
5. Факторы и уровни.

##### **Тема 4. Графики в R**

1. Типичные функции, используемые с факторами.
2. Графики в R

##### **Тема 5. Инструменты работы с Big Data**

1. Определения Big Data

##### **4.3 Оценочные средства для промежуточной аттестации.**

#### 4.3.1. Перечень компетенций образовательной программы. Индикаторы и критерии оценивания компетенций

Код компетенции	Наименование компетенции	Код индикатора компетенции	Наименование индикатора компетенций
СК ОС LA-25	Способен осуществлять подготовку, анализ, интерпретацию данных из первичных и вторичных источников, а также визуализировать и представлять результаты исследования	СК ОС LA-25.1	Решает задачи, связанные с коммуникацией компании с потребителями в сети интернет, в сфере её маркетинговой деятельности, используя теоретические знания и практические навыки

Индикатор оценивания	Критерии Оценивания
СК ОС LA-25.1 Решает задачи, связанные с коммуникацией компании с потребителями в сети интернет, в сфере её маркетинговой деятельности, используя теоретические знания и практические навыки	Базовый уровень - знает принципы сбора и анализа данных, преимущества и ограничения первичных и вторичных данных, базовые операции, применимые к различным типам данных. Повышенный уровень – может применять различные методы анализа данных и формулировать модели объяснения и интерпретации данных, критически описывать противоречия в данных.

#### 4.3.2 Типовые оценочные средства

Выбор вопросов для ответа на экзамене осуществляется случайным образом методом выбора студентом двух случайных номеров из общего списка. Каждому номеру соответствует определенный вопрос. Выбранные вопросы могут относиться к любой из изученных тем.

##### Примерные вопросы к экзамену

1. Классификации данных. Типы инструментов для работы с данными.
2. Основные концепции организации реляционных баз данных и хранилища данных (DWH)
3. Отличие баз данных от хранилищ данных (Digital WareHouses - DWH).
4. Сырые (raw) и нормализованные данные.
5. Основные принципы работы с R
6. Объектно-ориентированное программирование.
7. Векторы.
8. Матрицы и массивы.
9. Кадры данных.
10. Арифметические и логические операторы и значения.
11. Факторы и уровни.
12. Типичные функции, используемые с факторами.
13. Графики в R
14. Инструменты работы с Big Data

15. Определения Big Data
16. Принципы и подходы распределенной обработки больших данных.
17. Особенности HDFS.
18. Модель распределенных вычислений MapReduce.
19. Экосистема Hadoop. Фреймворк Apache Spark и его отличия от других баз данных.
20. Распределенная обработка данных и управление распределенными данными.

### Шкала оценивания

Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
Устный ответ на экзамене <sup>1</sup>	<ul style="list-style-type: none"> <li>Студент подробно излагает содержание вопроса: исчерпывающе, последовательно, четко и аргументированно излагает материал.</li> <li>Демонстрирует критическую оценку возможностей и ограничений представленного подхода/метода: рассматривает возможную критику, условия, в которых данные концепции или методы неприменимы.</li> <li>Демонстрирует способность проследить и реконструировать аргументацию авторов по теме, ссылается в ответе на первоисточники или актуальные исследования.</li> <li>Поясняет утверждение на уместных примерах.</li> <li>Уверенно отвечает на дополнительные вопросы, свободно ориентируется в теме.</li> </ul>	Отлично (41-60)
	<ul style="list-style-type: none"> <li>Студент подробно излагает содержание билета, но упускает некоторые аспекты рассматриваемого подхода/метода.</li> <li>Демонстрирует критическую оценку возможностей представленного подхода/метода, но не всегда способен проследить его ограничения.</li> <li>Демонстрирует способность частично реконструировать аргументацию авторов, на которых ссылается.</li> <li>Поясняет утверждение на уместных примерах.</li> <li>Отвечает на дополнительные вопросы по теме с небольшими паузами в разговоре.</li> </ul>	Хорошо (21-41)
	<ul style="list-style-type: none"> <li>Студент излагает содержание билета поверхностно;</li> <li>Демонстрирует критическую оценку возможностей представленного подхода/метода на уровне «здравого смысла».</li> <li>Демонстрирует способность частично реконструировать аргументацию авторов, на которых ссылается, но допускает ошибки.</li> <li>Поясняет утверждение на примерах, но не всегда сразу может обосновать их уместность.</li> <li>Отвечает на дополнительные вопросы по теме с паузами в</li> </ul>	Удовлетворительно (1-21)

<sup>1</sup> Преподавателям предлагается оценить ответ по 5 критериям, выставив за каждый критерий до 20 баллов, баллы выставляются по двум вопросам отдельно.



Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
	<p>разговоре и ошибками.</p> <ul style="list-style-type: none"> <li>• Студент излагает содержание билета сжато, не отражая сути вопроса;</li> <li>• Не демонстрирует критическую оценку возможностей представленного подхода/метода.</li> <li>• Не способен реконструировать аргументацию авторов, допускает серьезные ошибки.</li> <li>• Не использует примеры, либо предложенные примеры не отражают суть вопроса.</li> <li>• Не отвечает на дополнительные вопросы по теме.</li> </ul>	Неудовлетворительно (0)

**Итоговая семестровая оценка выставляется, исходя из формулы:  
40% текущий контроль 60% экзамен.**

#### **4.4. Методические материалы**

Текущий контроль осуществляется с целью мониторинга актуальной ситуации подготовки студентов к этапу промежуточного контроля. Текущий контроль осуществляется в формах: контроля посещаемости, ответов на практических занятиях, дискуссиях. Преподаватель отслеживает текущую ситуацию по указанным формам текущего контроля и выносит студентам рекомендации для улучшения качества подготовки к промежуточному контролю.

Промежуточный контроль по курсу осуществляется в форме экзамена.

Экзамен проводится в форме устного ответа включает ответ на 2 вопроса по билету. Преподавателям предлагается оценить ответ по 5-ти критериям (см.п.4.3.2), выставляя за каждый критерий до 20 баллов. Экзамен считается сданным, если студент прошел порог 41 балл, что соответствует оценке «удовлетворительно» и переносится в зачетную книжку и ведомость.

### **5. Методические указания для обучающихся по освоению дисциплины**

Дисциплина подразумевает работу над двумя взаимодополняющими видами деятельности: освоение содержания дисциплины в рамках практических занятий (а также самостоятельной работы) и решения практических задач.

Технология организации самостоятельной работы обучающихся включает использование информационных и материально-технических ресурсов образовательного учреждения.

Перед выполнением обучающимися внеаудиторной самостоятельной работы преподаватель проводит консультирование по выполнению задания, который включает цель задания, его содержания, сроки выполнения, ориентировочный объем работы, основные требования к результатам работы, критерии оценки. Во время выполнения обучающимися внеаудиторной самостоятельной работы и при необходимости преподаватель может проводить индивидуальные и групповые консультации.

Самостоятельная работа может осуществляться индивидуально или группами обучающихся в зависимости от цели, объекта, конкретной тематики самостоятельной работы, уровня сложности, уровня умений обучающихся.

**6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине**

**6.1. Основная литература.**

1. Миркин Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/450262>
2. Мхитарян В. С. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/450166>
3. Яковлев В. Б. Статистика. Расчеты в Microsoft Excel : учебное пособие для вузов / В. Б. Яковлев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 353 с. — (Высшее образование). — ISBN 978-5-534-01672-7. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/453051>

**6.2. Дополнительная литература.**

1. Истомина А. П. Анализ данных качественных исследований : практикум / составители А. П. Истомина. — Ставрополь : Северо-Кавказский федеральный университет, 2016. — 94 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/66014.html>
2. Добрина О. А. Анализ данных в социологии : учебное пособие / О. А. Добрина. — Новосибирск : Новосибирский государственный архитектурно-строительный университет (Сибстрин), ЭБС АСВ, 2013. — 101 с. — ISBN 978-5-7795-0666-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/68743.html>

**6.3. Учебно-методическое обеспечение самостоятельной работы.**

1. Положение об организации самостоятельной работы студентов федерального государственного бюджетного образовательного учреждения высшего образования «Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации» (в ред. приказа РАНХиГС от 11.05.2016 г. № 01-2211). [http://www.ranepa.ru/images/docs/prikazy-ranhigs/Pologenie\\_o\\_samostoyatelnoi\\_rabote.pdf](http://www.ranepa.ru/images/docs/prikazy-ranhigs/Pologenie_o_samostoyatelnoi_rabote.pdf)

**6.4 Нормативные правовые документы.**

1. Этический кодекс социолога: [https://www.ssa-rss.ru/index.php?page\\_id=84](https://www.ssa-rss.ru/index.php?page_id=84)
2. Профессиональный кодекс социолога: [https://www.ssa-rss.ru/index.php?page\\_id=73](https://www.ssa-rss.ru/index.php?page_id=73)

## 6.5. Интернет-ресурсы, справочные системы.

Электронные архивы отечественных журналов по социологии (выборочный список):

- Журнал социологии и социальной антропологии <http://www.jourssa.ru/>
- Социологический журнал [http://www.isras.ru/index.php?page\\_id=33](http://www.isras.ru/index.php?page_id=33)
- Социологическое обозрение <http://sociologica.hse.ru/>
- Социальная реальность [http://socreal.fom.ru/?link=PAST\\_RELEASES](http://socreal.fom.ru/?link=PAST_RELEASES)
- Социологические исследования <http://www.isras.ru/socis.html>

Базы данных, информационно-справочные и поисковые системы:

- портал «Экономика. Социология. Менеджмент» [www.ecsocman.ru](http://www.ecsocman.ru)
- веб-адреса научно-исследовательских и аналитических центров:
  - ФОМ – Фонд «Общественное мнение» <http://www.fom.ru/>
  - ВЦИОМ – Всероссийский Центр изучения общественного мнения <http://wciom.ru/>
  - Левада-центр [www.levada.ru](http://www.levada.ru)
  - ЦСПГИ – Центр социальной политики и гендерных исследований [www.socpolicy.ru](http://www.socpolicy.ru)

## 6.6. Иные рекомендуемые источники.

1. Christopher M. Bishop. (n.d.). Australian National University Pattern Recognition and Machine Learning.
2. Hastie, T., Tibshirani, R., Friedman, J. The elements of statistical learning: Data Mining, Inference, and Prediction. – Springer, 2009. – 745 pp.
3. Matthes, E. Python crash course: a hands-on, project-based introduction to programming. – No Starch Press, 2015. – 562 pp.
4. Гудфеллоу Я., Бенджио И., Курвилль А. - Глубокое обучение - Издательство "ДМК Пресс" - 2018 - 652с. - ISBN: 978-5-97060-618-6 - Текст электронный // ЭБС ЛАНЬ
5. Теория вероятностей и математическая статистика. Оценка параметров распределений : учебное пособие, Иванов, А. В., 2009
6. Nelli, F. (2018). Python Data Analytics : With Pandas, NumPy, and Matplotlib (Vol. Second edition). New York, NY: Apress.
7. Анализ данных в MS Excel : основные сведения о MS Excel, статистические таблицы и графики, статистические функции, пакет анализа (анализ данных) : учеб. пособие для вузов, Мхитарян, В. С., Шишов, В. Ф., 2018
8. Комбинаторика, Виленкин, Н. Я., Виленкин, А. Н., 2013
9. Несен А.В. - Microsoft Word 2010: от новичка к профессионалу - Издательство "ДМК Пресс" - 2011 - 448с. - ISBN: 978-5-94074-713-0 - Текст электронный // ЭБС ЛАНЬ

**7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы**  
*Требования к аудиториям (помещениям) для проведения занятий:*

Компьютерный класс, оснащенный компьютерами, объединенными в локальную вычислительную сеть с выходом в Интернет через шлюз. Презентационное оборудование (мультимедийный проектор, интерактивные доски и т.д.)

*Требования к программному обеспечению общего пользования:*

Установленное программное обеспечение MS Excel, R и R Studio