

**Санкт-Петербургский филиал федерального государственного
автономного образовательного учреждения высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"**

Факультет Санкт-Петербургская школа экономики и менеджмента
Национального исследовательского университета
«Высшая школа экономики»

Департамент прикладной математики и бизнес-информатики

**Рабочая программа дисциплины
Базы данных и аналитические системы**

для образовательной программы «Анализ больших данных в бизнесе, экономике и обществе»
направления подготовки 01.04.02 «Прикладная математика и информатика»
уровень магистратура

Разработчики программы
Мусабилов И.Л., преподаватель, ilya@musabirov.info
А.В.Сироткин, к.ф.-м.н., avsirotkin@hse.ru

Утверждена Академическим руководителем образовательной программы
«31» августа 2017 г.

А.В. Сироткин _____

Санкт-Петербург, 2017

*Настоящая программа не может быть использована другими подразделениями
университета и другими вузами без разрешения кафедры-разработчика программы.*



1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину «Базы данных и аналитические системы», учебных ассистентов и студентов направления подготовки 01.04.02 «Прикладная математика и информатика», обучающихся по образовательной программе «Анализ больших данных в бизнесе, экономике и обществе».

Рабочая программа дисциплины разработана в соответствии с:

- Образовательным стандартом НИУ ВШЭ по направлению подготовки 01.04.02 «Прикладная математика и информатика» (уровень магистратура), утвержденным ученым советом Национального исследовательского университета «Высшая школа экономики», протокол от 06.12.2013 г. № 50.
<https://www.hse.ru/data/2016/11/02/1111123560/01.04.02%20%D0%9F%D1%80%D0%B8%D0%BA%D0%BB%D0%B0%D0%B4%D0%BD%D0%B0%D1%8F%20%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0%20%D0%B8%20%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0.pdf>
- Образовательной программой «Анализ больших данных в бизнесе, экономике и обществе», направление подготовки 01.04.02 «Прикладная математика и информатика»;
- Объединенным учебным планом университета по образовательной программ «Анализ больших данных в бизнесе, экономике и обществе».

2 Цели освоения дисциплины

Целями освоения дисциплины «Базы данных и аналитические системы» является формирование навыков работы с анализом данных как процессом. Изучение основных структур и форм хранения данных.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент осваивает следующие компетенции:

Компетенция	Код по ОС ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции	Форма контроля уровня сформированности компетенции
<i>Системные компетенции</i>					
Способен принимать управленческие решения и готов нести за них ответственность	СК-5	РБ	Способен выбирать направление проекта с учетом имеющихся ресурсов, предлагать путь достижения цели и следовать ему.	Групповые выступления на семинарах, подготовка командного проекта	Домашнее задание



		СД	Знает основные системы хранения данных и способен выбирать наилучшие для проекта.		
<i>Профессиональные компетенции</i>					
<i>А) Социально-личностные компетенции</i>					
Способен к осознанному выбору стратегий межличностного взаимодействия	ПК-5	СД	Способен эффективно работать в команде, совместно добиваясь решения поставленной задачи.	Групповые выступления на семинарах, подготовка командного проекта	Домашнее задание.
<i>Б) Инструментальные компетенции</i>					
Способен описывать проблемы и ситуации профессиональной деятельности, используя язык и аппарат прикладной математики при решении междисциплинарных проблем.	ПК-14	СД	Способен формально описывать задачи, возникающие в бизнес-аналитике, и сводить их к математическим или технологическим задачам.	Семинарские занятия, подготовка командного проекта	Домашнее задание, доклад на семинаре, экзамен
Способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативных документов в профессиональной деятельности.	ПК-15	СД	Способен подготовить и согласовать ТЗ на разрабатываемый совместно программный продукт/модель. Способен сам следовать согласованному ТЗ и следить за выполнением ТЗ другими участниками.	Семинарские занятия, подготовка командного проекта	Домашнее задание
Способен строить и решать математические модели в соответствии с направлением подготовки и специализации	ПК-17	СД	Умеет сводить прикладные задачи к задачам регрессии, классификации и задачам предиктивного моделирования.	Лекции, семинарские занятия, подготовка командного проекта	Домашнее задание, экзамен

4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к циклу дисциплин по выбору студента магистерской программы «Анализ больших данных в бизнесе, экономике и обществе».



Изучение данной дисциплины базируется на следующих дисциплинах магистерской программы: «Алгоритмы и структуры данных», «Математические основы анализа данных»
Для освоения учебной дисциплины, студенты должны владеть следующими знаниями и компетенциями

- способен принимать управленческие решения и готов нести за них ответственность;
- способен к осознанному выбору стратегий межличностного взаимодействия;
- способен описывать проблемы и ситуации профессиональной деятельности, используя язык и аппарат прикладной математики при решении междисциплинарных проблем;
- способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативных документов в профессиональной деятельности
- способен строить и решать математические модели в соответствии с направлением подготовки и специализации.

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

- «Распределенная обработка и анализ больших данных»

5 Тематический план учебной дисциплины

Курс рассчитан на 32 часа аудиторной нагрузки, из них 16 часов лекций и 16 часов семинаров, общим объемом 4 зачетные единицы (152 часа).

№	Название раздела	Всего часов	Аудиторные часы		Самостоятельная работа
			Лекции	Семинары	
1	Введение в аналитику данных	34	4	4	26
2	Инфраструктура анализа данных	72	6	6	60
3	Организация аналитики в компании	46	6	6	34
ИТОГО		152	16	16	120

6 Содержание дисциплины

Раздел 1. Введение в аналитику данных

Аналитическое мышление. Бизнес-проблемы и наука о данных. Введение в предиктивное моделирование. Дата-продукты.

Раздел 2. Инфраструктура анализа данных.

Базы данных и хранилища данных. Функциональные классы аналитических систем. Системы оптимизации. Экспертные системы. Системы машинного обучения. Операционная бизнес-аналитика. Аналитическая отчетность. ERP-системы. Облачные решения анализ данных.

Раздел 3. Организация аналитики в компании.

Проектная и процессная организация аналитики. Business Intelligence. Business Analytics. Enterprise Decision Management. Data Science и Big Data.



7 Оценочные средства

7.1. Формы контроля знаний студентов

Тип контроля	Форма контроля	4 Модуль			Параметры
		апрель	май	июнь	
Текущий	Домашнее задание			*	Студенты в группах по 3-4 человека готовят приложение или сервис, основанный на данных
	Доклад на семинаре	*	*	*	Студенты в группах по 2-3 человека готовят 20-минутный доклад по одной из тем курса, согласованной с преподавателем.
Итоговый	Экзамен			*	Письменный экзамен - 60 минут

7.2. Критерии и шкалы оценки, примеры заданий

7.2.1. Текущий контроль

Домашнее задание

Примеры проектов для домашнего задания:

1. Сервис рекомендации музыкальных композиций, на основе акаунта из социальной сети.
2. Сервис отслеживания наиболее быстро распространяющихся новостей и прогноза их дальнейшей динамики.

Критерии оценки домашнего задания:

Домашнее задание ($O_{дз}$) проводится в виде сдачи программы или сервиса на платформе R/Shiny/ShinyApps, языке Python, платформе Azure ML или другой платформе, по согласованию с преподавателем. Работа оценивается “зачтено”/ “не зачтено”, при этом в случае оценки “зачтено” в формулу оценки в качестве балла за домашнюю работу проставляется 10, в противном случае -- 0.

Критериями оценки являются:

- корректность применения алгоритмов машинного обучения
- корректность определения прикладного программного интерфейса (API) и пользовательского интерфейса
- оригинальность идеи

Доклад на семинаре

Темы для доклада на семинаре:

1. Бизнес-проблемы и наука о данных. Формулировка бизнес-проблем. Решения, основанные на данных.
2. Проектная и процессная организация аналитики.
3. Business Intelligence. Business Analytics. Определение и соотношение понятий



4. Enterprise Decision Management. Суть концепции. Смысл управления решениями в организации
5. Data Science и Big Data. Определение и соотношение понятий с точки зрения бизнеса и инженерии
6. Базы данных и хранилища данных.
7. Функциональные классы аналитических систем.
8. Системы оптимизации. Экспертные системы. Системы машинного обучения.
9. Операционная бизнес-аналитика. Аналитическая отчетность. ERP-системы.
10. Облачные решения в области бизнес-аналитики
11. Инфраструктура анализа данных
12. Приложения и сервисы, основанные на данных
13. A/B тестирование и оптимизационные алгоритмы.

Критерии оценки доклада на семинаре:

- Полнота охвата выбранной темы
- Привлечение дополнительных источников литературы, в том числе из статей в ведущих журналах по теме, докладов ведущих конференций
- Структурирование презентации и управление временем доклада
- Уместное применение иллюстративного материала
- Корректность оформления ссылок
- Управление обсуждением доклада

Активное и конструктивное участие в обсуждении чужих докладов оценивается до 2 баллов к оценке за доклад на семинаре.

Максимальная оценка за доклад на семинаре – **10 баллов.**

7.2.2. Итоговый контроль

Экзамен

Примерные темы экзаменационных билетов:

1. Бизнес-проблемы и наука о данных. Формулировка бизнес-проблем. Решения, основанные на данных.
2. Проектная и процессная организация аналитики.
3. Business Intelligence. Business Analytics. Определение и соотношение понятий
4. Enterprise Decision Management. Суть концепции. Смысл управления решениями в организации
5. Data Science и Big Data. Определение и соотношение понятий с точки зрения бизнеса и инженерии
6. Базы данных и хранилища данных.
7. Функциональные классы аналитических систем.
8. Системы оптимизации. Экспертные системы. Системы машинного обучения.
9. Операционная бизнес-аналитика. Аналитическая отчетность. ERP-системы.
10. Облачные решения в области бизнес-аналитики
11. Инфраструктура анализа данных
12. Приложения и сервисы, основанные на данных



13. А/В тестирование и оптимизационные алгоритмы.

Критерии оценки за экзамен:

Экзамен состоит из набора закрытых и открытых вопросов, освещающих основные темы курса. Экзамен проводится в письменной форме. Максимальные баллы за каждый вопрос указываются в экзаменационном задании.

Максимальная оценка за экзамен – **10 баллов**.

7.3. Порядок формирования оценок по дисциплине

Накопленная оценка по дисциплине рассчитывается с помощью взвешенной суммы оценок за отдельные формы текущего контроля знаний следующим образом:

$$O_{\text{накопленная}} = 0,7O_{\text{дз}} + 0,3O_{\text{сем}}, \text{ где}$$

$O_{\text{дз}}$ - оценка знаний студента за домашнее задание;

$O_{\text{сем}}$ – оценка знаний студента за доклад на семинаре;

Результирующая оценка по дисциплине (которая идет в диплом) рассчитывается следующим образом:

$$O_{\text{результ}} = 0,8O_{\text{накопленная}} + 0,2O_{\text{экз}}, \text{ где}$$

$O_{\text{накопленная}}$ - накопленная оценка по дисциплине;

$O_{\text{экз}}$ - оценка за экзамен.

В формулу для $O_{\text{результ}}$ подставляются значения $O_{\text{накопленная}}$ и $O_{\text{экз}}$, округленные до ближайшего целого значения. $O_{\text{результ}}$ округляется до ближайшего целого значения.

По усмотрению ведущего преподавателя, если это не противоречит действующим документам на момент экзамена, при получении накопленной оценки 8 баллов и более, студент может быть освобожден от экзамена. В таком случае, с согласия студента, ему выставляется результирующая оценка, равная накопленной.

Студент не получает возможность пересдать низкие результаты за домашнюю работу а также при пропуске соответствующих им учебных часов.

При получении неудовлетворительной оценки $O_{\text{результ}}$ (значение после округления менее 4 баллов) выставляется оценка «НЕУДОВЛЕТВОРИТЕЛЬНО».

8 Образовательные технологии

Основными образовательными технологиями являются: интерактивные лекции, работа в группах на практических занятиях и семинарах.

9 Учебно-методическое и информационное обеспечение дисциплины

9.1 Основная литература

1. Cavanillas, J. New Horizons for a Data-Driven Economy [Electronic Resource] / José María Cavanillas, Edward Curry, Wolfgang Wahlster - Springer, 2016. – 303 p. - Authorized access: <http://link.springer.com/10.1007/978-3-319-21569-3> (Online Digital Library "Springer eBooks").



9.2 Дополнительная литература

1. Hastie T. The Elements of Statistical Learning: Data Mining, Inference, and Prediction [Electronic Resource] / Trevor Hastie, Robert Tibshirani, Jerome Friedman. - Springer, 2009.- Authorized access: <http://link.springer.com/book/10.1007/978-0-387-84858-7> (Online Digital Library "Springer eBooks").
2. Siegel E. Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die. Predictive Analytics [Electronic Resource] / Eric Siegel – Hoboken, N.J: Wiley, 2013. – 320 p. - Authorized access: <http://library.books24x7.com/toc.aspx?bookid=52840> (Online Digital Library "Books24x7").
3. Taylor J. Decision Management Systems: A Practical Guide to Using Business Rules and Predictive Analytics. Decision Management System [Electronic Resource] /James Taylor – NJ: IBM Press, 2012.- 316 p. - Authorized access: <https://library.books24x7.com/toc.aspx?bookid=45443> (Online Digital Library "Books24x7").
4. Blattberg R. Database Marketing: Analyzing and Managing Customers [Electronic Resource] / Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin New York, NY: Springer, - 2008, 871 p. (International Series in Quantitative Marketing Vol. 18). - Authorized access: <https://link.springer.com/content/pdf/10.1007%2F978-0-387-72579-6.pdf> (Online Digital Library "Springer eBooks").

10 Рекомендации для самостоятельной работы студентов.

Самостоятельная работа может рассматриваться как организационная форма обучения - система педагогических условий, обеспечивающих управление учебной деятельностью по освоению знаний и умений в области учебной деятельности без посторонней помощи. Студенту нужно четко понимать, что самостоятельная работа – не просто обязательное, а необходимое условие для получения знаний по дисциплине и развитию компетенций, необходимых в будущей профессиональной деятельности.

Самостоятельная работа проводится с целью:

- систематизации и закрепления полученных на лекциях теоретических знаний;
- углубления и расширения теоретических знаний;
- формирования умений использовать нормативную, правовую, справочную документацию и специальную литературу;
- развития познавательных способностей и активности студентов: творческой инициативы, самостоятельности, ответственности и организованности;
- формирования самостоятельности мышления, способностей к саморазвитию, самосовершенствованию и самореализации;
- формирования практических (общеучебных и профессиональных) умений и навыков;
- развития исследовательских умений;
- получения навыков эффективной самостоятельной профессиональной (практической и научно-теоретической) деятельности.

В учебном процессе выделяют два вида самостоятельной работы:

- аудиторная;
- внеаудиторная.

Аудиторная самостоятельная работа по дисциплине выполняется на учебных занятиях под непосредственным руководством преподавателя и по его заданию.

Внеаудиторная самостоятельная работа - планируемая учебная работа студентов, выполняемая во внеаудиторное время по заданию и при методическом руководстве преподавателя,



но без его непосредственного участия.

Самостоятельная работа, не предусмотренная программой учебной дисциплины, раскрывающей и конкретизирующей ее содержание, осуществляется студентом инициативно, с целью реализации собственных учебных и научных интересов.

Для более эффективного выполнения самостоятельной работы по дисциплине преподаватель рекомендует источники для работы, характеризует наиболее рациональную методику самостоятельной работы, демонстрирует ранее выполненные студентами работы и т. п.

Виды заданий для внеаудиторной самостоятельной работы, их содержание и характер могут иметь вариативный и дифференцированный характер, учитывать индивидуальные особенности студента.

Самостоятельная работа может осуществляться индивидуально или группами студентов online и на занятиях в зависимости от цели, объема, конкретной тематики самостоятельной работы, уровня сложности.

Контроль результатов внеаудиторной самостоятельной работы осуществляется в пределах времени, отведенного на обязательные учебные занятия по дисциплине на практических занятиях.

Для представления результатов самостоятельной работы рекомендуется:

Составить план выступления, в котором отразить тему, самостоятельный характер проделанной работы, главные выводы и/или предложения, их краткое обоснование и практическое значение – с тем, чтобы в течение 3 – 5 минут представить достоинства выполненного самостоятельно задания.

Подготовить иллюстративный материал в виде презентации для использования во время представления результатов самостоятельной работы в аудитории. Конкретный вариант наглядного представления результатов определяется форматом аудиторного занятия и задания преподавателя.

11 Материально-техническое обеспечение дисциплины и информационные технологии, используемые при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения информационных справочных систем (при необходимости).

Для проведения всех занятий используется проектор и компьютер для проекции слайдов.

12 Особенности организации обучения для лиц с ограниченными возможностями здоровья

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей, в том числе с применением электронного обучения и дистанционных технологий:

1) для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации.

2) для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации.



3) для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.