

Дополнительная профессиональная программа повышения квалификации «Основы хранилищ данных»

Программа направлена на формирование у слушателей базовых знаний и практических навыков по обработке, хранению и оптимизации данных с использованием интегрированной среды обработки.

№	Наименование	Характеристика
1	Наименование и цели программы	
1.1	Наименование	«Основы хранилищ данных»
1.2	УИН программы	ПК26-017
1.3	Вид программы	Дополнительная профессиональная программа повышения квалификации.
1.4	Цель	Научить проектировать хранилища данных, оптимизировать их работу и использовать для целей анализа данных.
1.5	Планируемые результаты (формируемые или совершенствующиеся компетенции в соответствии с указанным профстандартом)	Управление доступом к БД. Мониторинг событий, возникающих в процессе функционирования БД. Повышение производительности БД путем оптимизации выполнения запросов к БД. Профстандарт 06.011 «Администратор баз данных» (утв. приказом Министерства труда и социальной защиты Российской Федерации от 27 апреля 2023 г. № 408н).
2	Формат реализации	
2.1	Язык обучения	Русский
2.2	Форма обучения	Очно-заочная с применением дистанционных образовательных технологий (ДОТ)
2.3	Объем программы	40 академических часов
2.4	Режим занятий	В течение 6,5 недель в удобное для слушателя время по 3 - 9 ак. часов в неделю. Вебинары – согласно расписанию занятий.
2.5	Выдаваемый документ	Удостоверение о повышении квалификации Академии ФСБ.
2.6	Целевая аудитория	Специалисты по анализу данных; руководители проектов и менеджеры в области информационных технологий и

		аналитики; студенты и начинающие специалисты по анализу данных; лица, желающие приобрести базовые знания и навыки в области хранилищ данных.
2.7	Численность обучаемых по программе	Численность слушателей определяется договором об образовании, заключаемым при приеме на обучение.
3	Организационно-педагогические условия	
3.1	Учебный план	см. Приложение № 1
3.2	Календарный учебный график	см. Приложение № 2
3.3	Рабочая программа (учебные предметы, курсы, дисциплины (модули))	см. Приложение № 3
3.4	Профессорско-преподавательский состав ¹	Самородов Федор Анатольевич , преподаватель АНО ДПО «Академия ПСБ», преподаватель ОЧУ ДПО «Бауманский компьютерный учебный центр «Специалист.Ру», Агаджанян Артем Суренович , преподаватель АНО ДПО «Академия ПСБ», Колесникова Светлана Васильевна , старший преподаватель АНО ДПО «Академия ПСБ», Хомутичникова Ксения Сергеевна , преподаватель АНО ДПО «Академия ПСБ»
3.5	Форма итоговой аттестации	Экзамен в форме онлайн-тестирования.
3.6	Оценочные материалы	Пример задания см. Приложение № 4 . Тест содержит 25 вопросов, с двумя типами вопросов: с выбором одного ответа, с выбором нескольких ответов. За верный ответ на тестовый вопрос слушатель получает 2 балла. При частично правильном ответе на вопрос с выбором нескольких вариантов ответа слушатель получает 1 балл. При неверном выборе ответа баллы не начисляются. Максимальный балл, который может набрать слушатель за выполнение итогового тестирования – 50 баллов. Условием прохождения итоговой аттестации и получения удостоверения о повышении квалификации является прохождение итогового экзамена с результатом не менее 30 баллов.

¹ Сведения о преподавателях размещаются на сайте Академии ПСБ по адресу: <https://psb-academy.ru/organization/20>

3.7	Методические материалы	Учебная литература (учебники, пособия, книги) – предоставляется доступ к электронной библиотечной системе «Знаниум» (www.znanium.ru).
4	Контактная информация	
4.1	Руководитель программы	Мубаракшина Алсу Ирековна, директор Центра развития обучения по управлению данными АНО ДПО «Академия ПСБ» a.mubarakshina@psb-academy.ru

УЧЕБНЫЙ ПЛАН

дополнительной профессиональной программы повышения квалификации «Основы хранилищ данных»

Количество часов – 40 академических часов.

Форма обучения – очно-заочная с применением дистанционных образовательных технологий.

№ п/п	Наименование раздела/ модуля, темы	Общая трудоемкость, час.	Аудиторная работа, час.					С применением дистанционных образовательных технологий, электронного обучения, час.					Самостоятельная работа, час.	Текущий контроль успеваемости ²	Промежуточная аттестация (форма/час.) ³	Итоговая аттестация (вид/час.) ⁴	Код компетенции
			Всего	В форме практической подготовки	В том числе			Всего	В форме практической подготовки	В том числе							
					Лекции/в интерактивной форме	Практические (семинарские) занятия/в интерактивной форме	Контактная самостоятельная работа			Лекции/в интерактивной форме*	Практические (семинарские) занятия/в интерактивной форме	Контактная самостоятельная работа					
1	Введение	7,3						3,8		3,8			1,5	1,5**	0,5		A/03.4
2	Внутреннее устройство хранилищ данных	4,9						2,9		2,9			1	1			A/03.4
3	ETL для управления данными	6,5						4	2	2	2***		1,5	1			A/06.4

² Вид текущего контроля успеваемости: ПР – практическая работа.

³ Вид промежуточной аттестации: Т – тестирование

⁴ Вид итоговой аттестации: Э – экзамен.

* лекции в интерактивной форме отсутствуют.

** из них 0,5 ак.ч. – входное тестирование.

*** в интерактивной форме (вебинары).

4	Управление данными – Data Governance	4,8						1,8		1,8			1,5	1	0,5		A/06.4
5	Архитектура хранилища	5						3		3			1	1			A/03.4
6	Системы-потребители данных хранилища	5,3						2,8		2,8			1	1	0,5		A/03.4, B/03.5
7	Использование современного хранилища данных	5,2						4,4	2	2,4	2***		0,8				B/03.5
8	Итоговая аттестация	1														1	A/03.4, A/06.4, B/03.5
ИТОГО:		40						22,7	4	18,7	4		8,3	6,5	1,5	1	

*** в интерактивной форме (вебинары).

КАЛЕНДАРНЫЙ УЧЕБНЫЙ ГРАФИК

Период обучения:		
40 ак. часов	6,5 недель	1,5 месяца

РАБОЧИЕ ПРОГРАММЫ РАЗДЕЛОВ / МОДУЛЕЙ / ДИСЦИПЛИН

1. ВВЕДЕНИЕ
Тема 1.1. Входное тестирование.
Тема 1.2. Введение.
Тема 1.3. Предварительные требования.
Тема 1.4. Учебные базы данных.
Тема 1.5. Постановка задачи.
Тема 1.6. Структура запроса.
Тема 1.7. Операционные базы.
Тема 1.8. Подключение к серверу.
Тема 1.9. Хранилище данных.
Тема 1.10. Пример хранилища данных.
Тема 1.11. Практика.
2. ВНУТРЕННЕЕ УСТРОЙСТВО ХРАНИЛИЩ ДАННЫХ
Тема 2.1. Как устроены хранилища.
Тема 2.2. Факты и справочники.
Тема 2.3. Звезда или снежинка.
Тема 2.4. Календарь.
Тема 2.5. Медленно меняющаяся размерность (SCD).
Тема 2.6. Исторические таблицы.
Тема 2.7. Мусорная размерность (Junk Dimension).
Тема 2.8. Факты.
Тема 2.9. Быстро меняющаяся размерность.
Тема 2.10. Секционирование фактов.
Тема 2.11. Грануляция фактов.
Тема 2.12. Аддитивность мер.
Тема 2.13. Типы фактических таблиц.
3. ETL ДЛЯ УПРАВЛЕНИЯ ДАННЫМИ
Тема 3.1. ETL для управления данными.
Тема 3.2. Разбиение на шаги.
Тема 3.3. Перезапуск с любой операции.
Тема 3.4. Повторяемость запуска конвейера.
Тема 3.5. Практика.
4. УПРАВЛЕНИЕ ДАННЫМИ – DATA GOVERNANCE
Тема 4.1. Data Governance.
Тема 4.2. Качество данных.
Тема 4.3. Нормативные справочники.
Тема 4.4. Метаданные.
Тема 4.5. Практика.
5. АРХИТЕКТУРА ХРАНИЛИЩА
Тема 5.1. Архитектура хранилища.
Тема 5.2. Витрины данных.
Тема 5.3. Кимбал или Инмон.
Тема 5.4. Озеро данных (Data Lake).
Тема 5.5. Сложные типы хранилищ данных.

Тема 5.6. Слоеная архитектура хранилища.
6. СИСТЕМЫ-ПОТРЕБИТЕЛИ ДАННЫХ ХРАНИЛИЩА
Тема 6.1. Сценарии потребления данных.
Тема 6.2. Моделирование.
Тема 6.3. Многомерные модели данных.
Тема 6.4. Табличные модели данных.
7. ИСПОЛЬЗОВАНИЕ СОВРЕМЕННОГО ХРАНИЛИЩА ДАННЫХ
Тема 7.1. Современное хранилище данных.
Тема 7.2. Жизненный цикл данных.

Пример задания итоговой аттестации

Перечень примерных вопросов итогового экзамена с вариантами ответов:

1. Приведите разумные аргументы против использования хранилища данных. Почему использовать для аналитики операционные источники лучше, чем задействовать хранилище?
 - Используем знакомый и универсальный язык SQL;
 - **Не усложняется инфраструктура хранения данных (правильный ответ);**
 - **Высокая актуальность отчётов (правильный ответ);**
 - Можно задействовать несколько источников данных одновременно.
2. Для построения быстро меняющейся размерности (FCD) в хранилище необходимо добавить:
 - Столбцы к таблице размерности (справочнику);
 - **Столбцы к таблице фактов (правильный ответ);**
 - Таблицы, связанные с фактами;
 - Таблицы, связанные со справочниками (размерностями).
3. Какие данные имеет смысл хранить в Data Lake?
 - Справочные таблицы;
 - Фактические таблицы;
 - **Файлы с редко используемыми данными (правильный ответ);**
 - **Неструктурированные данные (правильный ответ).**
4. Какими качествами должно обладать хранилище данных?
 - **Единая база данных для всех источников (правильный ответ);**
 - На каждый источник данных требуется отдельное хранилище;
 - **Хранить данные вечно (правильный ответ);**
 - Хранить данные в исходном формате.
5. На каком этапе работают механизмы Data Quality?
 - **В ETL-процессе (правильный ответ);**
 - В модели данных;
 - В витрине данных;
 - В отчёте.