



III межвузовская конференция

**Бизнес-аналитика.**  
**Использование аналитической**  
**платформы Deductor в учебном**  
**процессе вуза**

28 июня 2016 года, г. Москва

# Академическая программа BaseGroup Labs

О результатах работы за 2011-2016 гг.

Николай Паклин



**BaseGroup Labs**  
ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ

## Статистика

- **104** вуза-партнера (2011 г. – **90**)
- **95** зарегистрированных преподавателей на сайте



## Результаты

- Летние и зимние школы (2012-2013 г.);
- Аудит и опрос всех вузов-партнеров;
- Новая редакция академической программы;
- Запуск портала преподавателей;
- Открытие электронного учебного центра на новой платформе;
- Обновленный электронный курс;
- Главы по бизнес-аналитике и Deductor в учебниках издательства «Юрайт».

# Летние и зимние школы

- Зимняя школа по анализу данных (2013 год, Н. Новгород)
- Летняя школа по анализу данных (2013 год, С.-Петербург)



## Новая редакция академической программы

- все существенные условия вынесены в «Положение об академической программы»;
- закреплён регламент взаимодействий с BaseGroup Labs и порядок предоставления ежегодных отчетов;
- введена возможность выдачи совместных сертификатов для студентов.

# Портал преподавателей



## Портал преподавателей

Уже **1 0 4** вуза-партнёра



**Навигатор**  
Все важное  
об академической программе



**Документы**  
Образцы соглашений  
и регламенты



**Материалы**  
Методические материалы  
и учебные пособия

### Блог: обучение бизнес-аналитике



18 мая 2016

**О доступе в электронный учебный центр**



5 апреля 2016

**Определение события в скоринговых моделях**



14 марта 2016

**О летней конференции преподавателей**



## Форум преподавателей

Задать вопрос

Сортировка

**Как пройти обучение в Электронном учебном центре** **3** КОММЕНТА  
20 мая 2016

Добрый день! Работаю доцентом в Брянском государственном техническом университете. Преподаю курс Интеллектуальный анализ данных в котором исполняется платформа Deductor. В Положении "Об академической программе BaseGroup Labs" сказано п. 2.7.1. "Неотъемлемой частью Программы является прохождение..."

 LagerevDG  
12 марта 2016

**Партнёрские отношения с НИУ ВШЭ - Санкт-Петербург** **1** КОММЕНТА  
21 февраля 2016

Добрый день! Я в настоящее время прорабатываю возможность заключения партнерского соглашения с кампусом НИУ ВШЭ - Санкт-Петербург (У Вас иду как преподаватель СПб Государственного Экономического Университета) Несколько вопросов: 1. НИУ ВШЭ - Москва является Вашим партнером - нужно ли новое...

 noskova  
20 февраля 2016

**Получение сертификата на основе разработанных учебных пособий** **1** КОММЕНТА  
30 сентября

Подписка на материалы сайта

Подписаться

# Электронный учебный центр



**BaseGroup Labs**  
ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ

Слушатель Тестовый Слушатель 3 ▾

Главная Мероприятия Уведомления Сообщения Ресурсы Форум F.A.Q. Профиль

Главная страница

Рекомендуемые мероприятия

№	Название	Тип	Вид	Начало	Окончание	Регистрация
1	<a href="#">Основы бизнес-аналитики и науки о данных (ЭКЗАМЕН)</a>	Дистанционное		2016-08-19 12:00	2016-08-19 20:00	Не начата

Ближайшие мероприятия

№	Название	Вид	Начало	Окончание	Операции
1	<a href="#">Основы бизнес-аналитики и науки о данных (ЭКЗАМЕН)</a>		19.08.2016 г. О 12:00	19.08.2016 г. О 20:00	
2	<a href="#">Разработка скоринговых моделей</a>		17.10.2016 г. О 09:30	20.10.2016 г. О 14:00	

Объявления

14 июня 2016 г. О 00:00 [Открытие центра](#)

13 июня 2016 г. О 00:00 [ЦЕНТР ЭЛЕКТРОННОГО УЧЕБНОГО ЦЕНТРА "Старого" п](#)

[Все объявления](#)

	Лекция 8. Системы управления мастер-данными		<p><b>Баллы</b></p> <p>Статус: <b>Пройден</b></p> <p>Первый запуск: 2016-06-07 18:03</p> <p>Последний запуск: 2016-06-07 18:03</p> <p>Затрачено времени: 0 ч 0 мин</p> <p>Попыток: 1 из Неограничено</p>	0%	<a href="#">Запустить</a>
	Лекция 9. Преобразование данных. Часть 1		<p><b>Баллы</b></p> <p>Статус: <b>Пройден</b></p> <p>Первый запуск: 2016-06-07 18:03</p> <p>Последний запуск: 2016-06-07 18:03</p> <p>Затрачено времени: 0 ч 0 мин</p> <p>Попыток: 1 из Неограничено</p>	0%	<a href="#">Запустить</a>
	Лекция 10. Преобразование данных. Часть 2		<p><b>Баллы</b></p> <p>Статус: <b>Пройден</b></p> <p>Первый запуск: 2016-06-07 18:03</p> <p>Последний запуск: 2016-06-07 18:03</p> <p>Затрачено времени: 0 ч 0 мин</p> <p>Попыток: 1 из Неограничено</p>	0%	<a href="#">Запустить</a>
	Самопроверочный тест (21)		<p><b>76 Баллы</b></p> <p>Статус: <b>Пройден</b></p> <p>Первый запуск: 2016-06-06 08:45</p> <p>Последний запуск: 2016-06-08 14:35</p> <p>Затрачено времени: 6 ч 15 мин</p> <p>Попыток: 2 из 10</p>	76%	75 (баллов) <a href="#">Запустить</a>



**BaseGroup Labs**  
ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ

## Введение в бизнес-аналитику и науку о данных

Лекция 7. Большие данные. Наука о данных




### Централизованное ХД с витринами данных

ЦХД является основным местом хранения исторических данных, используемых для целей бизнес-аналитики. Однако оно не решает всех задач, поскольку:

- по мере того, как ЦХД развивается, обеспечить доступ к данным становится все труднее, усложняется администрирование и обеспечение информационной безопасности;
- ЦХД является инструментом масштаба предприятия, поэтому не может содержать данные в виде, удобном для работы всеми подразделениями, по всем направлениям бизнеса;
- хранилище используется одновременно многими пользователями, которые конкурируют между собой за доступ к данным внутри ЦХД, вызывая конфликты и нагружая хранилище;
- коллективный доступ большого числа пользователей может привести к искажению информации в ЦХД, ее частичной или даже полной потере;
- недостаточная пропускная способность и неадекватность телекоммуникационных линий в случае территориально распределенной компании.



Все эти и многие другие проблемы, связанные с эксплуатацией ЦХД, позволяют снять использование *витрин данных*.

### Лекция 7. Базовые архитектуры КИФ

**Какие из целей будущего проекта по бизнес-аналитике реальны, и будут четко понятны руководству?**

- выявить факторы, влияющие на отток абонентов

скоринговую карту  
уровень одобрений по кредитным заявкам, сохранив текущий список в портфеле  
ые совместно покупаемые товары  
точности прогноза продаж в 99,9%  
вывать абонентов

Пропустить Ответить

Это свейгар. Он представляет собой последовательность операций с данными, представленную в виде иерархического дерева. В дереве каждая операция образует узел. Свейгар состоит из ветвей, каждая из которых начинается с узла импорта данных.

Параметр	Значение
Время выполнения	00:00:00
Процесс остановлен по условию останова	False
Процесс остановлен в результате останова	False
Имя файла	True
Процесс остановлен пользователем	True
Имя файла	C:\Базовые шаблоны работы в Deductor Studio\Примеры.tst
Экспортировать все поля	True
Отправить на e-mail	False
Дополнительная работа	-
Разрешить загрузку данных	-
Формат даты	dd/MM/yyyy
Разрешить загрузку временных	True
Формат времени	hh:mm
Отправлять строки	True
План	True
Источники	False
Путь	True
Первая строка является заголовком	True

# Основы бизнес-аналитики и науки о данных

Стартовый курс по бизнес-аналитике: введение в анализ, интеграция и подготовка данных, аудит выборок, визуализация

Введение в  
бизнес-  
аналитику и  
науку о данных

Подготовка  
данных

Основы витрин и  
хранилищ  
данных

Визуализация  
данных

Аудит и  
профайлинг  
данных

Существенная переработка материалов

# Введение в бизнес-аналитику и науку о данных

- От разведочного анализа Тьюки к Data Mining
- Методология CRISP-DM
- Большие данные – основные понятия
- Наука о данных (Data Science) – история и современное понимание термина
- Специалист по данным и бизнес-аналитик: отличия

## Подготовка данных

- Виды источников данных
- Интеграция данных и ее методы
- Корпоративная информационная фабрика и ее компоненты
- Процессы и архитектуры КИФ
- Системы управления мастер-данными
- Преобразование данных

# Основы витрин и хранилищ данных

- Модели MOLAP и ROLAP
- Плотность заполнения гиперкуба в MOLAP
- Витрины данных
- Историчность в хранилище данных
- Медленно меняющиеся измерения

# Аудит и профайлинг данных

- Понятия аудита и профайлинга
- Механизмы формирования пропусков: MCAR, MAR, MNAR
- Алгоритмы и методы обработки пропусков в зависимости от их механизма формирования
- Модели выбросов и экстремальных значений

# Использование материалов BGL

1. Любой контент электронных курсов и книг, издаваемых при поддержке BaseGroup, является авторской разработкой компании и ее сотрудников.
2. Компания BaseGroup Labs **не поддерживает** инициативы преподавателей по использованию ее контента в издаваемых преподавателями методических пособиях, УМК, монографиях и т.д. в том виде, который признается специализированными информационными системами поиска плагиата как «наличие заимствований».
3. Компания BaseGroup Labs **одобрительно относится к** инициативам преподавателей по разработке преподавателями собственного контента на основе знаний, умений и навыков, полученных из электронных курсов и книг BaseGroup Labs.

# Книги и учебники



2014 г.



Август 2016 г.