

Паспорт  
оценочных материалов для проведения текущего контроля и  
промежуточной аттестации обучающихся по дисциплине (модулю)  
Прикладные задачи больших данных

Перечень оценочных материалов и индикаторов достижения компетенций, сформированность которых они контролируют

Наименование оценочного средства	Коды индикаторов достижения формируемых компетенции	Номер приложения
Выполнение и собеседование по лабораторным работам	ИД-1 ОПК-5 ИД-2 ОПК-5 ИД-1 ПК-1 ИД-2 ПК-1 ИД-1 ОПКИ-6 ИД-2 ОПКИ-6 ИД-3 ОПКИ-6	1
Экзамен	ИД-1 ОПК-5 ИД-2 ОПК-5 ИД-1 ПК-1 ИД-2 ПК-1 ИД-1 ОПКИ-6 ИД-2 ОПКИ-6 ИД-3 ОПКИ-6	2

Утверждено на заседании кафедры «Информационные системы»  
протокол № 3 от « 11 » октября 2021 года  
Заведующий кафедрой \_\_\_\_\_ А.А. Романов

## I. Текущий контроль

## Приложение 1

### Выполнение и собеседование по лабораторным работам

#### 1. Процедура выполнения лабораторных работ и собеседования по ним

Количество проводимых лабораторных работ в течение всего периода освоения дисциплины	7 работ
Формат представления результатов	Бумажный отчет
Общее количество вопросов для собеседования	4-10 вопросов
Количество основных задаваемых при собеседовании вопросов	3-4 вопроса
Методические рекомендации (при необходимости)	Выдаются индивидуально перед началом выполнения работы

#### 2. Шкала оценивания с учетом срока сдачи

Критерии оценивания	Балл
Студент правильно выполнил задание работы, продемонстрировал знания теоретического и практического материала	Отлично
Студент правильно выполнил задание работы, продемонстрировал неполные знания теоретического и практического материала	Хорошо
Студент выполнил задание работы, но допустил значительные неточности при выполнении, продемонстрировал неполные знания теоретического и практического материала	Удовлетворительно
Студент неправильно выполнил задание работы, не продемонстрировал знания теоретического и практического материала	Неудовлетворительно

#### 3. Перечень лабораторных работ и вопросов при собеседовании

#### **Лабораторная работа к разделу 1: Использование инструментов Data Mining Client для Excel 2007 в подготовке данных**

Данная лабораторная работа описывает возможности инструментов, относящихся к Data Mining Client для Excel 2007, в части подготовки данных для анализа.

Microsoft Excel имеет большое число статистических функций. Некоторые являются встроенными, некоторые доступны после установки пакета анализа. Для выполнения лабораторной работы используется именно это программное обеспечение. Средства, включенные в пакет анализа данных, доступны через команду Анализ данных меню Сервис, в т.ч. используя пакет клиент интеллектуального анализа данных (DataMiningClient), который также входит в набор надстроек интеллектуального анализа.

Для выполнения лабораторной работы необходимо/желательно сделать полную установку надстроек, в которую входит и DataMiningClient.

**Цель работы:** определить наличие связи между признаками, используя набор данных.

Задание включает следующие этапы:

- анализ предметной области;
- постановка задачи;
- подготовка данных;
- построение моделей;
- проверка и оценка моделей;
- выбор модели;
- применение модели;
- коррекция и обновление модели.

Примерные вопросы при собеседовании:

1. Назовите основную цель анализа данных?
2. Приведите наиболее распространенные определения информации.
3. Приведите формализованное описание факта.
4. Что называется фактографическими данными?
5. Перечислите основные составляющие экспертной системы?
6. Назовите основные задачи, решаемые средствами ИАД?
7. Перечислите принципы анализа данных, предложенные Дж. Тьюки?
8. Назовите три уровня анализа информации. Чем они отличаются? 9. Чем отличается классификация от кластеризации?
9. Назовите различие между данными и знаниями.

**Лабораторная работа к разделу 2:** Использование аналитической платформы «Дедуктор» в подготовке данных

1. Создайте новый проект и сохраните его под именем проба1.ded.
2. Заполните свойства проекта.
3. Просмотрите файл проекта через любой текстовый редактор.
4. Сделайте видимой вкладку Подключения.
5. Создайте текстовый файл с данными, содержащими буквенные, цифровые форматы, а также формат даты, в виде матрицы, и сохраните его в текстовом редакторе на рабочем диске под именем тест2.txt.
6. Создайте новый проект и сохраните его под именем проба2.ded.
7. Проведите импорт подготовленного текстового файла в созданный проект, настройте параметры импорта таким образом, чтобы данные файла отображались корректно.

8. Метку узла переименуйте посредством контекстного меню в Проба импорта. В комментарии к узлу впишите: Текстовый файл тест2.txt с разделителями.
9. Добавьте к узлу узел Настройка набора данных и задайте следующие метки к столбцам: Поле1, Поле2, Поле3 и далее по количеству столбцов.
10. Экпортируйте набор данных в текстовый файл с настройками, предлагаемыми по умолчанию, сохраняя файл с названием пробаэкспорт.txt на рабочем диске.
11. Импортируйте только что экспортированный файл в Deductor, задав в поле Мастера импорта путь сохраненного файла.
12. Присоедините к новому узлу импорта (путем копирования) предыдущую ветвь, начиная с узла Настройка набора данных.
13. Между экспортом и настройкой набора данных вставьте еще один узел настройки, в котором измените тип столбца Поле 3 на логический.
14. Удалите только что вставленный узел.
15. Сохраните проект.

#### **Примерные вопросы при собеседовании:**

1. Из каких частей состоит Deductor?
2. Какие варианты поставки Deductor существуют?
3. Чем отличается версия Professional от Academic?
4. Имеются ли ограничения по количеству обрабатываемых записей в версии Deductor Academic?
5. Сколько категорий пользователей Deductor можно выделить?
6. Перечислите функции аналитика.
7. Сколько вкладок на панели управления Deductor Studio?
8. Что такое проект в Deductor Studio?
9. Какое расширение имеет файл проекта?
10. Как создать новый проект?
11. Как сохранить текущий проект под другим именем?
12. Как отредактировать свойства проекта?
13. Сколько мастеров имеется в Deductor Studio?
14. Как скопировать ветвь сценария при помощи механизма drag & drop?
15. Какие шаги мастера импорта нужно пройти для импорта текстового файла? 1
6. Что позволяет сделать обработчик Настройка набора данных?

#### **Лабораторная работа к разделу 3: Статистический подход к анализу данных**

**Задание:** изучить взаимосвязь между двумя экономическими явлениями с помощью модели парной линейной регрессии:

1. Найти статистическую информацию о следующих показателях субъектов РФ по Центральному, Северо-Западному и Южному федеральным округам за 2018 год<sup>1</sup>:
  - **Валовой региональный продукт на душу населения, руб.**
  - **Фактическое конечное потребление домашних хозяйств на душу населения на территории субъектов РФ (в текущих рыночных ценах), руб.**
2. Сформировать таблицу исходных данных.
3. На основе теоретического анализа обосновать связь между изучаемыми признаками, выбор факторного и результативного признаков.
4. Построить поле корреляции и сформулировать гипотезу о наличии и форме связи между переменными модели.
5. Оценить параметры линейного уравнения парной регрессии, дать интерпретацию полученных оценок параметров.
6. Оценить качество полученной модели (коэффициент детерминации, статистическая значимость параметров регрессии и корреляции с помощью  $F$ -критерия и  $t$ -критерия, доверительные интервалы для параметров).
7. Выполнить прогноз среднедушевого фактического конечного потребления домашних хозяйств при прогнозном значении валового регионального продукта на душу населения, составляющем 110% от среднего уровня. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.
8. Составить аналитическое заключение.

### *Методические рекомендации*

- 1) Работа выполняется в MS Excel. Калькулятор не использовать. Числа вручную в ячейки и формулы не вносить.
- 2) При формировании выборки не учитывать данные для территорий «в том числе» — Ненецкого АО и Архангельской области без автономного округа. Должно получиться 36 наблюдений.
- 3) На листе 1 разместить таблицу с данными:

№ наблюдения	Регионы	ФАКТИЧЕСКОЕ КОНЕЧНОЕ ПОТРЕБЛЕНИЕ ДОМАШНИХ ХОЗЯЙСТВ на душу населения (в текущих рыночных ценах, рублей)	ВАЛОВОЙ РЕГИОНАЛЬНЫЙ ПРОДУКТ на душу населения (рублей)
1	1 Белгородская область	283 520,6	506 420,9
2	2 Брянская область	243 024,8	253 100,4
3	3 Владимирская область	243 681,8	300 273,6
4	4 Воронежская область	335 266,0	370 610,4
5	5 Ивановская область	221 052,1	182 398,1
6	6 Калужская область	278 622,9	411 565,3
7	7 Костромская область	231 282,9	256 848,4
8	8 Курская область	265 532,5	346 340,2
9	9 Липецкая область	305 869,7	431 821,2
10	10 Московская область	407 628,0	509 545,2
11	11 Орловская область	245 554,2	285 354,0
12	12 Рязанская область	244 572,1	320 764,0
13	13 Смоленская область	243 281,8	296 289,6
14	14 Тамбовская область	269 915,2	289 847,1

- 4) На этом же листе построить поле корреляции (предварительно выполнить п.3 задачи, значения регрессора расположить по оси абсцисс, значения объясняемой переменной — по оси ординат).
- 5) Расчеты оценок параметров и качества регрессии выполнить с использованием инструмента «Регрессия» пакета анализа Excel.
- 6) К п. 7 задания. Точечный прогноз заключается в получении прогнозного значения зависимой переменной  $\hat{y}_p$  подстановкой в уравнение регрессии прогнозируемого значения независимой переменной  $x_p$ :

$$\hat{y}_p = a + b \cdot x_p$$

<sup>1</sup> Сайт Росстата [www.gks.ru](http://www.gks.ru), раздел Публикации, подраздел Статистические издания, сборник «Регионы России. Социально-экономические показатели 2020» <https://rosstat.gov.ru/folder/210/document/13204>, раздел «Валовой региональный продукт»

*Интервальный прогноз* состоит в построении доверительного интервала, с заданной вероятностью содержащего точную величину прогнозируемого значения  $y_p$ . Рассчитывается стандартная ошибка прогноза:

$$m_{\hat{y}_p} = \sqrt{D(\varepsilon) \cdot \left( 1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)},$$

где  $D(\varepsilon) = \frac{\sum (y_i - \hat{y}_i)^2}{n-2}$ .

Вычисляется предельная ошибка прогноза, которая с вероятностью  $p = 1 - \alpha$  не будет превышена:

$$\Delta \hat{y}_p = t_{\text{таб}} \cdot m_{\hat{y}_p}$$

Границы доверительного интервала прогноза определяются по формуле:

$$y_p = \hat{y}_p \pm \Delta \hat{y}_p.$$

Таким образом, прогнозное значение с заданной вероятностью  $p$  будет находиться в границах:

$$\hat{y}_p - \Delta \hat{y}_p \leq y_p \leq \hat{y}_p + \Delta \hat{y}_p.$$

### Примерные вопросы при собеседовании:

1. На какие вопросы позволяет ответить визуальный анализ поля корреляции?
2. Назовите метод оценки параметров линейной регрессии, который Вы использовали в данной работе? На какой идее основан данный метод?
3. Как интерпретируются параметры линейной регрессии?
4. Перечислите критерии, которые используются для оценки качества полученной регрессионной модели.
5. Что показывает коэффициент детерминации  $R^2$ ?
6. Какова цель расчета F-статистики? В чем состоит нулевая гипотеза при проверке значимости уравнения регрессии в целом?
7. По каким параметрам определяется критическое значения F-статистики в процедуре оценки значимости уравнения в целом?
8. Как проверить статистическую значимость отдельных параметров регрессии?
9. Какова цель построения доверительных интервалов для параметров модели?
10. Как выполнить прогноз по уравнению регрессии?

### Лабораторная работа к разделу 4: Анализ последовательностей и прогнозирование данных

**Задание:** используя данные о численности населения в РФ, построить ряд динамики за последние 10 лет, для которого:

- 1) Рассчитать:
  - а) цепные и базисные показатели динамики: абсолютный прирост, темп роста, темп прироста, результаты расчетов представить в табличном виде;
  - б) среднегодовой уровень ряда динамики;
  - в) средний абсолютный прирост, средний темп роста, средний темп прироста;
- 2) Произвести:
  - а) сглаживание ряда динамики трехлетней скользящей средней.
  - б) аналитическое выравнивание ряда динамики по прямой, используя метод наименьших квадратов.
 Результаты расчетов представить в таблице.
- 3) Изобразить фактический и выровненные ряды графически.



$T_{\text{пр}}^{\text{ч}}\%$										
------------------------------	--	--	--	--	--	--	--	--	--	--

**Среднегодовой уровень ряда динамики ( $\bar{y}$ )** рассчитывается в зависимости от вида ряда и расстояния между его уровнями (для правильного выбора формулы расчета необходимо определить вид ряда динамики):

- для интервального ряда:

– с равноотстоящими уровнями — по средней арифметической простой:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n},$$

где  $n$  – число уровней;

– с неравноотстоящими уровнями — по средней арифметической взвешенной:

$$\bar{y} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i},$$

где  $t_i$  – периоды времени, при которых значение уровня не изменяется;

- для моментного ряда:

– с равноотстоящими уровнями — по средней хронологической простой:

$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1} = \frac{\frac{y_1 + y_n}{2} + \sum_{i=2}^{n-1} y_i}{n-1};$$

– с неравноотстоящими уровнями – по средней хронологической взвешенной:

$$\bar{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \dots + (y_{n-1} + y_n)t_{n-1}}{2 \sum_{i=1}^{n-1} t_i} = \frac{\sum_{i=1}^{n-1} (y_i + y_{i+1})t_i}{2 \sum_{i=1}^{n-1} t_i},$$

где  $t$  – длительность интервала между соседними уровнями.

**Средний абсолютный прирост ( $\bar{\Delta}$ )** за весь период, ограничивающий ряд динамики, определяется как средняя арифметическая простая из цепных абсолютных приростов:

$$\bar{\Delta} = \frac{\sum \Delta_{ц}}{n-1} = \frac{y_n - y_1}{n-1},$$

где  $n$  – число уровней ряда.

**Средний темп роста ( $\bar{T}_p$ )** выражается в процентах (или числом, как коэффициент роста  $\bar{K}_p$ ) и рассчитывается:

- для равноотстоящих рядов по средней геометрической из цепных коэффициентов роста:

$$\begin{aligned} \bar{T}_p &= \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \frac{y_4}{y_3} \cdot \dots \cdot \frac{y_n}{y_{n-1}}} \cdot 100\% = \sqrt[n-1]{K_{2/1} \cdot K_{3/2} \cdot \dots \cdot K_{n/n-1}} \cdot 100\% = \\ &= \sqrt[n-1]{\prod K_{ц}} \cdot 100\% \end{aligned}$$

или после сокращения:



$$\bar{T}_p = \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \frac{y_4}{y_3} \cdot \dots \cdot \frac{y_n}{y_{n-1}}} \cdot 100\% = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\%;$$

- для неравноотстоящих рядов определяется по средней геометрической, взвешенной по продолжительности периодов:

$$\bar{T}_p = \sqrt[\sum t_i]{\prod (K_{\text{п}})^{t_i}},$$

где  $t_i$  – интервал времени, в течение которого сохраняется данный темп роста.

**Средний темп прироста ( $\bar{T}_{\text{пр}}$ )** определяют на основании среднего темпа роста:

$$\bar{T}_{\text{пр}} = \bar{T}_p - 100\% \text{ - в процентах;}$$

$$\bar{K}_{\text{пр}} = \bar{K}_p - 1 \quad \text{— как коэффициент.}$$

2) Расчеты и результаты сглаживания ряда динамики трехлетней скользящей средней и методом аналитического выравнивания ряда динамики по прямой  $y_t = a_0 + a_1 t$  представить в таблице 4.2.

Сглаживание ряда динамики простой скользящей средней состоит в замене его исходных уровней средними арифметическими из соседних  $m$  уровней, которые располагаются по центру интервала сглаживания (в нашем случае  $m = 3$ ). Недостаток метода — теряются первые и последние уровни ряда.

Для аналитического выравнивания методом наименьших квадратов с целью упрощения расчетов показатели времени можно подобрать так, чтобы их сумма была равна нулю (см. гр.3 табл. 4.2):

$$\sum t = 0.$$

Тогда параметры уравнения тренда можно вычислить, используя формулы:

$$a_0 = \frac{\sum y}{n}, \quad a_1 = \frac{\sum ty}{\sum t^2}.$$

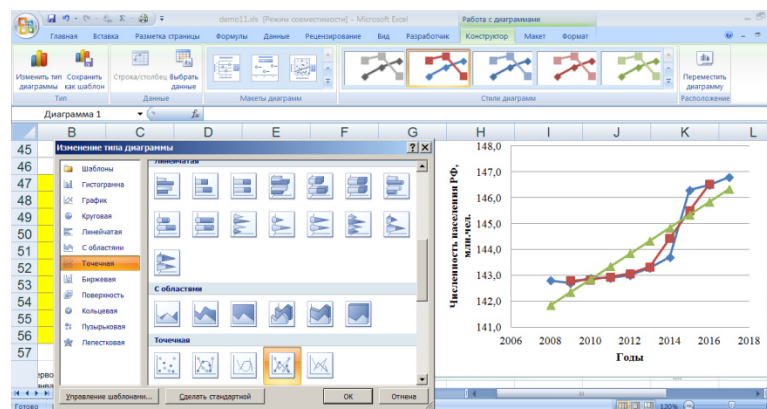
Расчетные значения  $\hat{y}$  получают подстановкой значений  $t$  в полученное уравнение тренда.

**Таблица - Выявление основной тенденции ряда динамики**

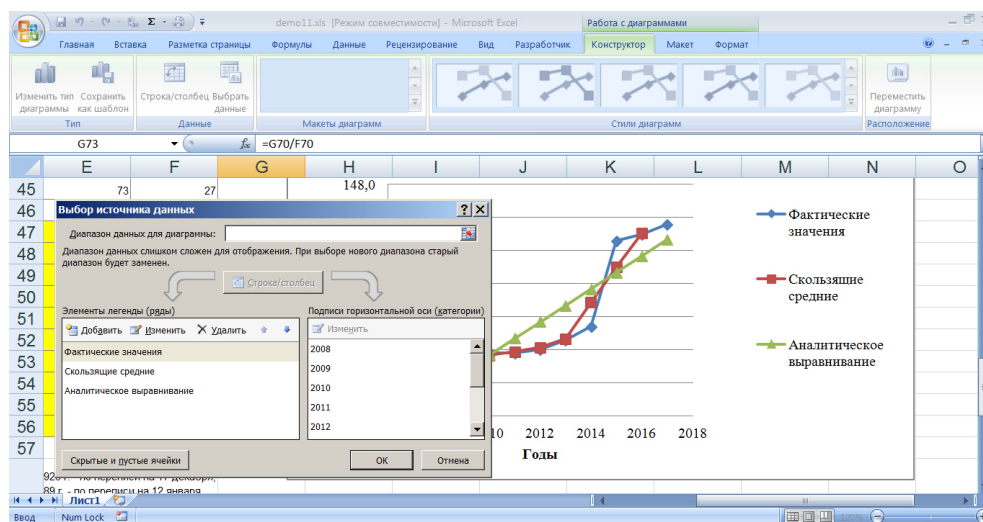
Годы	Наименование признака, ед. изм. ( $y_t$ )	Трехлетние скользящие средние	$t$	$t^2$	$ty$	$\hat{y}$ (аналитическое выравнивание)
А	1	2	3	4	5	6
20	...	—	-9	...	...	...
20	...	...	-7	...	...	...
...	...	...	-5	...	...	...
...	...	...	-3	...	...	...
...	...	...	-1	...	...	...
...	...	...	1	...	...	...
...	...	...	3	...	...	...

Годы	Наименование признака, ед. изм. ( $y_t$ )	Трехлетние скользящие средние	$t$	$t^2$	$ty$	 (аналитическое выравнивание)
...	...	...	5	...	...	...
20...	...	...	7	...	...	...
20...	...	—	9	...	...	...
Итого	...	X	0	...	...	...

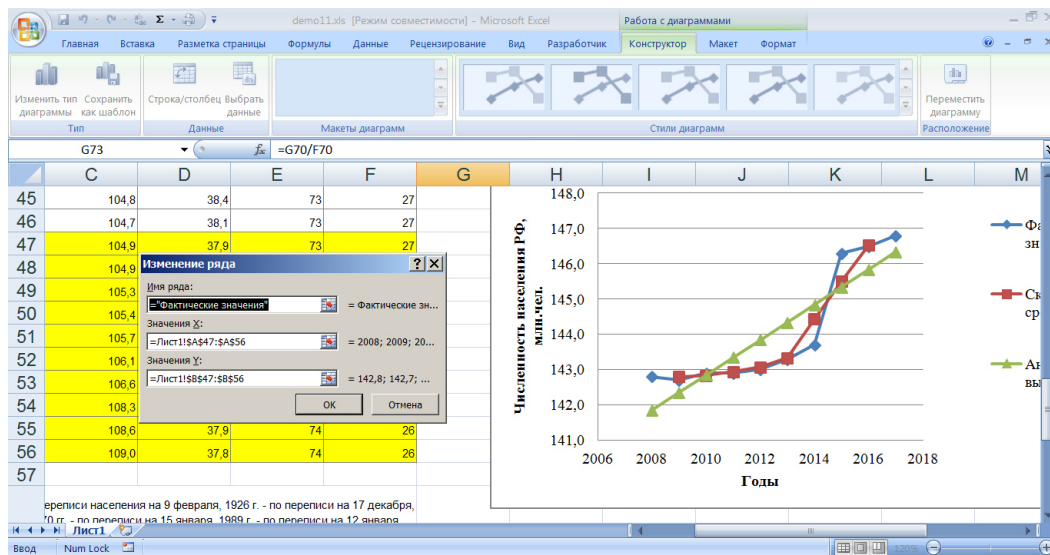
3) Для графического изображения наблюдаемых значений уровней ряда динамики и результатов выравнивания необходимо использовать мастер диаграмм Excel, шаблон «Точечная»:



При выборе данных при построении диаграммы необходимо получить три ряда с помощью команды «Выбор данных»:



При этом требуется указать соответствующие массивы значений для каждого из рядов:



— фактические значения — по оси  $x$  — массив значений гр.А, по оси  $y$  — массив значений гр.1 табл. 4.2;

— скользящие средние — по оси  $x$  — массив значений гр.А (первый и последний годы не включаются, т.к. соответствующих им данных не будет), по оси  $y$  — массив значений гр.2 табл.4.2;

— аналитическое выравнивание — по оси  $x$  — массив значений гр.А, по оси  $y$  — массив значений гр.6 табл.4.2.

Образец графического изображения наблюдаемых и выровненных данных представлен ниже.

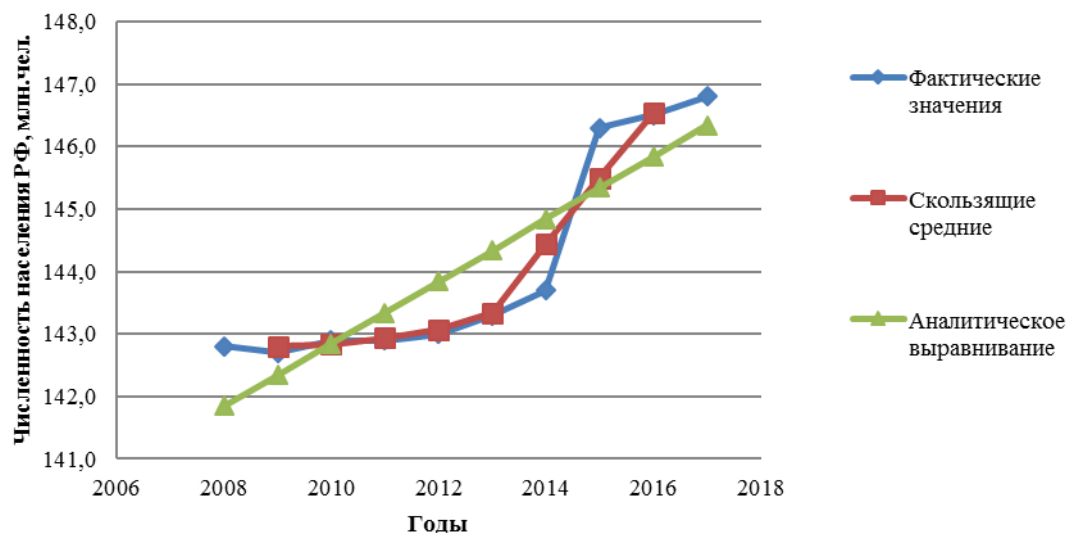


Рисунок — Фактический и выровненные ряды.

**Примерные вопросы при собеседовании:**

1. Какой тип данных называется рядом динамики? В чем состоит отличие рядов динамики от пространственных (кросс-секционных) выборок?
2. Перечислите относительные показатели, используемые для характеристики временного ряда.
3. В чем состоит различие между цепными и базисными показателями ряда динамики?
4. Может ли темп роста быть отрицательным числом?
5. Может ли темп прироста быть отрицательным числом?
6. Какова цель использования средних для характеристики динамического ряда?
7. От каких характеристик временного ряда зависит выбор формулы расчета среднего уровня ряда?
8. В чем заключается идея метода сглаживания ряда простой скользящей средней?
9. В чем заключается аналитическое выравнивание динамического ряда по прямой?
10. Назовите метод, который используется для оценки параметров линейного тренда?

### **Лабораторная работа к разделу 5: Кластерный анализ с использованием карт Кохонена**

Работа направлена на освоение основных методов и способов кластеризации с использованием самоорганизующихся карт Кохонена, освоение принципов построения и использования простейших нейронных сетей, приобретение практических навыков по использованию инструментария Deductor Studio.

Задание: имеется база данных коммерческих банков с показателями деятельности за текущий период. Необходимо провести их кластеризацию, т.е. выделить однородные группы банков на основе показателей из базы данных, всего показателей - 21. Исходная таблица находится в файле "banks.xls". Она содержит показатели деятельности коммерческих банков за отчетный период.

Основные этапы:

1. Разработать сценарии построения самоорганизующихся карт Кохонена.
2. Создать сводную таблицу банков, включив в нее суммарные сведения. Таблицу получить путем слияния соответствующих полей из разных таблиц и последующей группировки.
3. Для подготовленной сводной таблицы разработать сценарий кластеризации с использованием самоорганизующихся карт Кохонена.
4. Создать отчеты по всем разработанным сценариям.

### **Примерные вопросы при собеседовании:**

1. Поясните необходимость использования карт Кохонена при кластеризации.
2. Объясните общий принцип построения самоорганизующейся карты признаков Кохонена.
3. Каким образом производится назначение размеров и формы ячеек (нейронов) карты Кохонена?
4. Как осуществляется назначение начальных значений весовых коэффициентов нейронов?
5. Поясните понятия скорости и радиуса обучения нейросети.

## 6. Какие критерии используются для остановки процесса обучения?

**Лабораторная работа к разделу 6:** Разновидности и технический анализ графиков чартов; Основные фигуры технического анализа. Осцилляторы методы технического анализа.

Для выполнения лабораторной работы необходимо/желательно сделать полную установку надстроек, в которую входит и DataMiningClient для Excel 2007.

Средства, включенные в пакет анализа данных, доступны через команду Анализ данных меню Сервис, в т.ч. используя пакет клиент интеллектуального анализа данных (DataMiningClient), который также входит в набор надстроек интеллектуального анализа.

**Задание:** Данные в таблице показывают цены закрытия для акций Reuters за пятидесятидневный период.

Используя данные таблицы6

### 1. Построить скользящую среднюю.

Скользящая средняя – индикатор с задержкой во времени, который позволяет «сглаживать» данные и используется для подтверждения трендов цен. Период, выбранный для расчета скользящей средней, зависит от ценной бумаги, которую рассматривают с помощью этого метода; обычно используются периоды в 9, 10, 18, 20, 40, 50, 100 и 200. На фьючерсных рынках обычно используются краткосрочные средние, например с периодом 9 и 18, а для долгосрочных вложений, например в акции, чаще применяются периоды 50, 100 или 200. Если считать, что у данной бумаги наблюдается циклическая периодичность торговли, то по данным исследований наиболее удачным показателем будет скользящая средняя с периодом, связанным с торговым циклом. Широко используются три вида скользящих средних, каждая из которых имеет свои достоинства и недостатки, а именно:

- простая скользящая средняя (ПСС);
- взвешенная скользящая средняя (ВСС);
- экспоненциальная скользящая средняя (ЭСС).

### 2. Определить индикаторы движения:

- индекс относительной силы (RSI);
- стохастический осциллятор;
- схождение/расхождение скользящих средних (MACD).

Индекс относительной силы (RSI) был разработан Дж. Уэллсом Уайлдером-младшим первоначально для анализа графиков в виде «баров» акций или финансовых индексов, но сейчас применяется на всех рынках. Индикатор сравнивает эффективность ценной бумаги только с ее собственной эффективностью в прошлом, а не с другими бумагами или рынком в целом.

Стохастические осцилляторы впервые были разработаны как инженерный аналитический инструмент и были адаптированы для определения состояния перекупленности/перепроданности с использованием простой процентной шкалы американским аналитиком Джорджем С. Лэйном. Основное применение индикатора – поиск различий между стохастическими линиями и собственно ценой ценной бумаги. Эта

информация может использоваться для подтверждения торговых решений. Стохастические осцилляторы основаны на наблюдении за ценами:

- Когда цены на рынке во время тренда падают, цены закрытия группируются в самой нижней части ценового диапазона.

- Когда цены на рынке во время тренда поднимаются, цены закрытия группируются в самой верхней части ценового диапазона.

Стохастический анализ существует в быстрой и медленной модификациях. Быстрый осциллятор использует две линии, изображаемые разными цветами в программах или сплошной и пунктиром на иллюстрациях. Простое значение линии %K (сплошной) изображается на шкале от 0 до 100 %. Другая линия, называемая %D, представляет собой простую скользящую среднюю от %K (изображается пунктиром).

Схождение/расхождение скользящих средних (MACD): Джеральд Аппель разработал индикатор MACD для определения направления тренда и момента его изменения, первоначально при отслеживании 13– и 26-недельных циклов рынка. Индикатор использует две линии экспоненциальной скользящей средней, которые колеблются выше или ниже нулевой линии и дают сигналы перекупленности/перепроданности. В отличие от других индикаторов, в MACD отсутствуют верхняя и нижняя границы значения, например от 0 до 100.

#### **Примерные вопросы при собеседовании:**

1. Общая характеристика методов фильтрации;
2. Скользящие средние (Moving Averages) и их особенности;
3. Типы скользящих средних;
4. Сигналы, подаваемые скользящими средними;
5. Использование комбинаций скользящих средних;
6. Фильтры на скользящих средних;
7. Горизонтальное расположение точек скользящих средних;
8. Другие способы применения скользящих средних;
9. Индекс Относительной Силы (Relative Strength Index);
10. Стохастические линии (Stochastics) ;
11. Использование скользящих средних для создания осцилляторов;
12. Метод конвергенции-дивергенции (Moving Averages Convergence-Divergence);
13. Накопление Объема (Volume Accumulation);
14. Сводные рекомендации по использованию осцилляторов всех типов.

#### **Лабораторная работа к разделу 7: Социальная ответственность ПАО «Сбербанк»**

<b>Исходные данные к разделу «Социальная ответственность:»</b>	
1. Описание рабочего места (рабочей зоны, технологического процесса, используемого оборудования) на предмет возникновения: - вредных проявлений факторов	1. Анализ эффективности программ КСО на предприятии – ПАО «Сбербанк».

<p>производственной среды</p> <ul style="list-style-type: none"> <li>- (метеоусловия, вредные вещества, освещение, шумы, вибрация, электромагнитные поля, ионизирующие излучения);</li> <li>- опасных проявлений факторов производственной среды (механической природы, термического характера, электрической, пожарной природы);</li> <li>- негативного воздействия на окружающую природную среду (атмосферу, гидросферу, литосферу);</li> <li>- чрезвычайных ситуаций (техногенного, стихийного, экологического и социального характера).</li> </ul>	<p>ПАО «Сбербанк» – это крупнейший банк Российской Федерации</p>
<p>2. Список законодательных и нормативных документов по теме.</p>	<p>1. миссия Сбербанка России;</p> <p>2. Положения стратегия Сбербанка России;</p> <p>3. принципы, закреплённые международными стандартами в области КСО и устойчивого развития: ISO 26000, AA1000;</p> <p>4. Стандарты раскрытия информации, разработанные группой GRI (Global Reporting Initiative);</p> <p>5. Политика предприятия в области корпоративной и социальной ответственности.</p>
<p><b>Перечень вопросов, подлежащих исследованию, проектированию и разработке:</b></p>	
<p>1. Анализ факторов внутренней социальной ответственности:</p> <ul style="list-style-type: none"> <li>- принципы корпоративной культуры исследуемой организации;</li> <li>- системы организации труда и его безопасности;</li> <li>- развитие человеческих ресурсов через обучающие программы и программы</li> </ul>	<p>Провести анализ внутренней социальной ответственности организации:</p> <ul style="list-style-type: none"> <li>– безопасность труда;</li> <li>– стабильность выплаты заработной платы;</li> <li>– развитие человеческих ресурсов через обучающие программы подготовки и повышения квалификации;</li> <li>– системы медицинского и социального</li> </ul>

<p>подготовки и повышения квалификации;</p> <ul style="list-style-type: none"> <li>- системы социальных гарантий организации;</li> <li>- оказание помощи работникам в критических ситуациях.</li> </ul>	<p>страхования.</p>
<p>2. Анализ факторов внешней социальной ответственности:</p> <ul style="list-style-type: none"> <li>- содействие охране окружающей среды; - взаимодействие с местным сообществом и местной властью;</li> <li>- спонсорство и корпоративная благотворительность;</li> <li>- ответственность перед потребителями товаров и услуг (выпуск качественных товаров);</li> <li>-готовность участвовать в кризисных ситуациях и т.д.</li> </ul>	<p>Провести анализ внешней социальной ответственности компании;</p> <ul style="list-style-type: none"> <li>– анализ мероприятий по корпоративной благотворительности;</li> <li>– анализ мероприятий по социальной поддержке населения;</li> <li>– анализ мероприятий экологического влияния.</li> </ul>
<p>3. Правовые и организационные вопросы обеспечения социальной ответственности:</p> <ul style="list-style-type: none"> <li>- анализ правовых норм трудового законодательства;</li> <li>- анализ специальных (характерные для исследуемой области деятельности) правовых и нормативных законодательных актов;</li> <li>- анализ внутренних нормативных документов и регламентов организации в области исследуемой деятельности.</li> </ul>	<p>Провести анализ на основе:</p> <ul style="list-style-type: none"> <li>– правовых норм трудового законодательства внутренних нормативных документов и регламентов организации в области исследуемой деятельности.</li> </ul>
<p><b>Перечень графического материала:</b></p>	
<p>1. Эскизные графические материалы к расчётному заданию</p>	<p>1. Структура стейкхолдеров ПАО «Сбербанк»;</p> <p>2. Приоритеты Банка в области КСО;</p> <p>3. Структура программы КСО «Экономическое влияние»;</p> <p>4. Структура программы КСО «Социальное</p>



	<p>влияние»;</p> <p>5. Структура программы КСО «Экологическое влияние»;</p> <p>6. Затраты в области финансирования социально-значимых проектов в регионах присутствия группы Сбербанка;</p> <p>7. Затраты в области инвестиций в человеческий капитал;</p> <p>8. Затраты в области благотворительности и спонсорства</p>
--	--

### **Примерные вопросы при собеседовании:**

1. Поточковые данные, формирующие базу больших данных;
2. процедура структуризации и записи данных;
3. Функция обращения к данным;
4. Визуализация и первичный статистический анализ больших данных;
5. Финансовые технологии как новые участники рынка:
6. Тенденции развития финансового рынка:
7. Особенности правового регулирования FinTech:
8. Влияние FinTech на банковскую систему:
9. FinTech-стартапы и их привлекательность для инвесторов:
10. Главные FinTech тренды на финансовом рынке.

Экзамен

1. Процедура проведения

Общее количество вопросов к экзамену	24 вопроса
Количество вопросов в билете	2 вопроса
Наличие задач в билете	нет
Формат проведения	Устно

2. Шкала оценивания с учетом текущего контроля работы обучающегося в семестре

Критерии оценки уровня сформированности компетенций по дисциплине	Балл
выставляется обучающемуся, если он показал глубокие знания материала по поставленному вопросу, грамотно, логично и стройно его излагает	Отлично
выставляется обучающемуся, если он твердо знает материал, грамотно его излагает, но допускает несущественные неточности в ответе на вопрос	Хорошо
выставляется обучающемуся, если он показывает знания только основных положений по поставленному вопросу, требует в отдельных случаях наводящих вопросов для принятия правильного решения, допускает отдельные неточности	Удовлетворительно
выставляется обучающемуся, если он допускает грубые ошибки в ответе на поставленный вопрос	Неудовлетворительно

3. Примерный перечень вопросов к экзамену

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Виды массивов данных.
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R.
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов.
8. Методики анализа больших данных.
9. Процесс аналитики анализа больших данных.
10. Характеристика Big Data на мировом рынке.
11. Характеристика Big Data в России.
12. . Понятие Data Mining.
13. Вопросы безопасности больших данных.
14. Когнитивный анализ данных.
15. Модели данных.
16. Статистические методы анализа данных.

17. Различия между параметрическими, непараметрическими и номинальными методами.
18. Корреляционный анализ.
19. Регрессионный анализ.
20. Основная идея дисперсионного анализа.
21. Сущность кластерного анализа.
22. Дискриминантный анализ: модель и общая процедура выполнения.
23. Цели факторного анализа.
24. Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки.