

УДК 378

DOI: <https://doi.org/10.32820/2074-8922-2020-66-29-41>

ОБЗОР ПРОГРАММНЫХ ПРОДУКТОВ ДЛЯ РЕАЛИЗАЦИИ ДИДАКТИЧЕСКОЙ РЕДУКЦИИ: EXCEL, STATISTICA, SPSS

© Шматков Д.И.

Українська інженерно-педагогічна академія

Информация про автора:

Шматков Даниил Игоревич: ORCID: 0000-0003-2952-4070; d.shmatkov@uipa.edu.ua; кандидат педагогических наук, доцент, докторант кафедры креативной педагогики и интеллектуальной собственности; Украинская инженерно-педагогическая академия; ул. Университетская, 16, г. Харьков, 61003, Украина.

В статье представлен обзор таких статистических программных продуктов, как Excel, STATISTICA и SPSS, в контексте реализации разработанного и подтвердившего свою эффективность метода дидактической редукции. Использование разработанных критериев позволило обосновать эффективность применения программных продуктов на следующих этапах: формирование исходной совокупности объектов исследования и частных показателей, которые характеризуют определенные аспекты развития этих объектов; расчет интегрального показателя для получения обобщенной оценки уровня развития объектов исследования; идентификация влияния каждого частного показателя на изменения интегрального показателя и редукция показателей, осуществляющих наименьшее влияние; группировка объектов исследования и редукция части объектов, исходя из усредненных характеристик групп.

Система критериев оценки программных продуктов состояла из следующего: распространенность и доступность; гибкость и функциональность (возможность ввода формул вручную; возможность задания / изменения типа, вида, параметров иллюстраций; возможность реализовать метод в одном средстве); интерфейс (организация в виде каскадных окон; возможность получать результаты в отдельном файле / окне; представление доступных функций в иерархическом виде); скорость анализа (необходимость дополнительных действий для анализа; четкость, логичность и наглядность представления полученных результатов; общая производительность, учитывая затраты времени на анализ результатов); наличие дополнительных специальных функций.

Представлен детальный пример реализации метода дидактической редукции в процессе изучения предметной области развития социальных медиа и больших данных. В результате комбинирования Excel, STATISTICA и SPSS, используя три наиболее подходящих варианта комбинаций, установлено, что наиболее эффективным является объединение Excel и STATISTICA. При этом Excel предложено использовать на первом и втором, а STATISTICA на третьем и четвертом этапах реализации метода дидактической редукции.

Ключевые слова: дидактическая редукция; социальные медиа; большие данные; EXCEL; STATISTICA; SPSS.

Шматков Д.І. «Огляд програмних продуктів для реалізації дидактичної редукції: EXCEL, STATISTICA, SPSS»

У статті представлено огляд таких статистичних програмних продуктів, як Excel, STATISTICA і SPSS, в контексті реалізації розробленого методу дидактичної редукції. Використання розроблених критеріїв уможливило обґрунтування ефективності використання програмних продуктів на таких етапах: формування вихідної сукупності об'єктів дослідження і окремих показників, які характеризують певні аспекти розвитку цих об'єктів; розрахунок інтегрального показника для отримання узагальненої оцінки рівня розвитку об'єктів дослідження; ідентифікація впливу кожного окремого показника на зміни інтегрального показника і редукція показників, які здійснюють найменший вплив на інтегральний показник; групування об'єктів дослідження і редукція частини об'єктів, виходячи з усереднених характеристик груп.

Система критеріїв оцінки програмних продуктів складалася з: поширеність і доступність; гнучкість і функціональність (можливість введення формул вручну; можливість завдання / зміни типу, виду, параметрів ілюстрацій; можливість реалізувати метод в одному засобі); інтерфейс (організація у вигляді каскадних вікон; можливість отримувати результати в окремому файлі / вікні; представлення доступних функцій в ієрархічному вигляді); швидкість аналізу (необхідність додаткових дій для аналізу; чіткість, логічність і наочність представлення отриманих результатів; загальна продуктивність, враховуючи витрати часу на аналіз результатів); наявність додаткових спеціальних функцій.

Представлено детальний приклад реалізації методу дидактичної редукції в процесі навчання предметної галузі розвитку соціальних медіа та великих даних. У результаті комбінування Excel, STATISTICA і SPSS, використовуючи три найбільш адекватні варіанти комбінацій, встановлено, що найефективнішим є поєднання Excel і STATISTICA. При цьому Excel запропоновано використовувати на першому і другому, а STATISTICA на третьому і четвертому етапах реалізації методу дидактичної редукції.

Ключові слова: дидактична редукція; соціальні медіа; великі дані; EXCEL; STATISTICA; SPSS.

D. Shmatkov "A review of software products for implementing didactic reduction: EXCEL, STATISTICA, AND SPSS"

In the article, a review of statistical software products such as Excel, STATISTICA, and SPSS is provided in the context of implementation of a didactic reduction method which was developed and proved to be effective. Taking into account the developed criteria, the effectiveness of using software products at the following stages was substantiated: formation of the initial set of research objects and particular indicators that characterize certain aspects of the development of these objects; calculation of an integral indicator to obtain a generalized assessment of the level of development of the research objects; identification of the influence of each particular indicator on changes in the integral indicator and reduction of indicators that have the least impact; clustering of the research objects and reduction of part of the objects based on the averaged characteristics of the groups.

The system of criteria for evaluating software products consisted of the following: prevalence and availability; flexibility and functionality (the ability to enter formulas manually; the ability to set / change a type and parameters of illustrations; the ability to implement the method using a single tool); interface (the presence of cascading windows; the ability to receive results in a separate file / window; presentation of available functions in a hierarchical form); speed of analysis (the need for additional steps for analysis; clarity, consistency and visibility of results presentation; overall performance, taking into account the time spent on analysis); the presence of additional special functions.

A detailed example of the implementation of the didactic reduction method in the process of studying the subject area of the development of social media and big data is presented. As a result of combining Excel, STATISTICA and SPSS, using the three most suitable combinations, it was defined that combining Excel and STATISTICA is the most effective way for the method implementation. In this case, Excel is proposed to be used at the first and second stages, and STATISTICA – at the third and fourth stages.

Keywords: didactic reduction; social media; big data; EXCEL; STATISTICA; SPSS.

1. Введение.

Постановка проблемы. С развитием информационных технологий доступ к информации, а также ее объем неуклонно растет. Выбор авторитетных источников и снижение когнитивной сложности обучения в подобных условиях ложится на плечи преподавателей. Подобные процессы, как и другие интервенции в учебный процесс, должны содержать научно обоснованные подходы. Теоретическое обоснование, разработка и имплементация подобных подходов является актуальной проблемой.

Анализ последних исследований и публикаций. При подготовке учебного материала широко применяется метод (4S TMD), включающий этапы отбора, структурирования, описания и дидактической редукции [1]. В литературе методы, позволяющие упростить восприятие студентами содержания обучения, сохранив при этом основные посылы, принято объединять общим термином «дидактическая

редукция» [2], [3]. То есть дидактическая редукция является интегральным процессом, который происходит на всех этапах подготовки содержания обучения.

Программные продукты являются адекватным средством применения дидактической редукции, так как они вмещают в себя возможности упрощения сложной реальности, моделирования, сведения к качественным переменным, иллюстративного или символического представления и пр.

Дидактическую редукцию применяют в виде отражения предметов первой необходимости в качестве первоочередной задачи при разработке, например, веб-атласов [4]. Известным является программное средство обучения теории графов, направленное на понимание логической структуры многочисленного количества экспериментов [5]; ученые отмечают, что программное средство упрощает восприятие предметной области за счет ее рационального уменьшения.

В другой работе с целью уменьшения когнитивной нагрузки студентов предложено применение такого метода дидактической редукции, как использование визуального программирования в начале процесса изучения текстовых или профессиональных языков программирования [6]. Целесообразность применения дидактической редукции в данном контексте подтверждена и более ранними исследованиями [7].

Представленные научные труды посвящены реализации конкретных узких проблем обучения. Внедрение компьютерных средств обучения как общий метод дидактической редукции рассматривается в литературе не часто. Так, предложено применение программных разработок, обеспечивающих работу с таблицами, диаграммами, бизнес-ландшафтами и прочее, с целью оптимизации работы преподавателей в области формулировки, выбора и структурирования содержания и задач [8]. Применение дидактической редукции посредством подобных программ реализуется с целью обеспечения совместимости содержания обучения по времени с другими учебными событиями, то есть обеспечение устойчивости учебного прогресса. Авторами также подчеркнута склонность студентов к работе с задачами, которые имеют высокую степень структурирования, рекомендациями и инструкциями, которые включают в себя эскиз идеи для решения и, по меньшей мере, гарантии частичного успеха в решении.

Рассмотрим разработанный метод дидактической редукции, применимый к различным дисциплинам с метрологическим контекстом [3], [9] – [11]. Данный метод предусматривает реализацию последовательности следующих этапов: формирование исходной совокупности объектов исследования и частных показателей, которые характеризуют определенные аспекты развития этих объектов → расчет интегрального показателя для получения обобщенной оценки уровня развития объектов исследования → идентификация влияния каждого частного показателя на изменения интегрального показателя и редукция показателей, осуществляющих наименьшее влияние → группировка объектов исследования и редукция некоторых объектов, исходя из усредненных характеристик групп.

С применением объединенных математических инструментов преподаватель

может формировать программу курса или отдельные темы, согласовывая их с доступным временем на обучение. Студенты могут применять полученные знания в будущем в процессе выполнения профессиональных обязанностей. Реализация метода не является сложной или специфичной. Он может быть реализован с применением общедоступных программных средств в короткий промежуток времени. При этом выбор наиболее оптимальных для решения данной задачи программных средств является важным вопросом.

Целью представленной статьи является оценка и выбор по системе разработанных критериев программных средств наиболее адекватных для применения метода дидактической редукции содержания дисциплины метрологического направления.

2. Теоретические основы исследования.

В научных работах исследуются различные варианты использования средств программного обеспечения для математического моделирования и критерии их оценки. Известным является подход [12], при котором проводят оценку использования различных средств программного обеспечения для расчета линейной регрессии для одной переменной по таким критериям, как легкость, интуитивная понятность, способность выполнять необходимые функции. В других разработках [13] необходимо обратить внимание на такие критерии, как гибкость, наличие встроенных аналитических средств, удобство интерфейса, простота работы, производительность, популярность, простота освоения.

В данной статье система критериев была сформирована с учетом рассмотренных источников, принципов дидактики, а также системного анализа. Проблема решается в контексте разработки теоретико-методического инструментария учебной деятельности через внедрение междисциплинарного подхода к исследованиям, который позволяет повысить качество и обоснованность выводов и рекомендаций.

Таким образом, разработан набор критериев системы оценки средств программного обеспечения для реализации метода дидактической редукции, а именно:

- 1) распространенность и доступность;
- 2) гибкость и функциональность;
- 2.1) возможность ввода формул вручную;

- 2.2) возможность задания / изменения типа, вида, параметров иллюстраций;
- 2.3) возможность реализовать метод в одном средстве;
- 3) интерфейс;
- 3.1) организация в виде каскадных окон;
- 3.2) возможность получать результаты в отдельном файле / окне;
- 3.3) представление доступных функций в иерархическом виде;
- 4) скорость анализа;
- 4.1) необходимость дополнительных действий для анализа;
- 4.2) четкость, логичность и наглядность представления полученных результатов;
- 4.3) общая производительность, учитывая затраты времени на анализ результатов;
- 5) наличие дополнительных специальных функций.

3. Результаты исследования.

Учитывая доступный уровень техники и наиболее распространенные практики, реализацию метода дидактической редукции [9] наиболее целесообразно провести одним из трех вариантов использования средств программного обеспечения.

Вариант 1. Использование MS Excel на всех этапах.

Вариант 2. Использование MS Excel на первом и втором и STATISTICA на третьем и четвертом этапах.

Вариант 3. Использование MS Excel на первом и втором и SPSS на третьем и четвертом этапах.

Проведем оценку этих вариантов по системе критериев на примере применения метода для исследования в целях обучения развития социальных медиа и больших данных на предприятиях Украины, которые занимаются различными видами деятельности.

3.1. Использование MS Excel на всех этапах реализации метода.

Согласно варианту 1, реализация всех этапов метода проводится в программе Excel.

На первом этапе была сформирована исходная совокупность из 13 объектов исследования, в которую были включены предприятия разнообразных видов экономической деятельности Украины.

Выбор таких объектов объясняется тем, что развитие социальных медиа и больших данных в производственных целях зависит от специфики деятельности предприятия, для которого они формируются. А специфика деятельности определяет принадлежность предприятия к определенному виду деятельности. Как правило, принадлежность закреплена соответствующими государственными классификаторами. Поэтому в процессе исследования особенностей развития социальных медиа и больших данных в Украине (или других странах мира) целесообразно включить в исследование различные виды экономической деятельности.

Перечень частных показателей для оценки развития социальных медиа и больших данных на предприятиях Украины приведен в табл. 1.

Таблица 1

Частные показатели оценки развития социальных медиа и больших данных

Код	Название показателя	Тип влияния
i.1	Представление предприятия или рекламирование его работы (товаров, услуг) в социальных медиа	Стимулятор
i.2	Получение отзывов клиентов или предоставление ответов на их вопросы в социальных медиа	Стимулятор
i.3	Привлечение клиентов в развитие или инновацию товаров и услуг в социальных медиа	Стимулятор
i.4	Сотрудничество с деловыми партнерами или другими организациями в социальных медиа	Стимулятор
i.5	Блоги или микроблоги предприятия в социальных медиа	Стимулятор
i.6	Вебсайты с мультимедийным содержанием в социальных медиа	Стимулятор
i.7	Средства обмена знаниями в социальных медиа	Стимулятор
i.8	Развитие больших данных своего предприятия, полученных со смарт-устройств или датчиков геолокационных данных	Стимулятор
i.9	Развитие больших данных, полученных с портативных устройств	Стимулятор
i.10	Развитие больших данных, сформированных из социальных медиа	Стимулятор
i.11	Развитие больших данных, сформированных из других источников	Стимулятор

В таблиці 2 представлені исходні значення показателів і.1 – і.11 для підприємств різних видів діяльності в Україні в 2018 році.

Таблиця 2

Значення частних показателів оцінки розвитку соціальних медіа і великих даних на підприємствах різних видів діяльності [14]

Об'єкт	Показатель										
	i.1	i.2	i.3	i.4	i.5	i.6	i.7	i.8	i.9	i.10	i.11
В середньому по видах діяльності в Україні	940,3	688,3	435,1	702,3	286,4	510,3	465,3	212,3	177,1	138,7	149,0
Переробляюча промисловість	2795,0	1880,0	1187,0	2028,0	774,0	1499,0	1239,0	626,0	482,0	375,0	420,0
Постачка електроенергії, газу і кондиціонованого повітря	148,0	124,0	54,0	104,0	48,0	97,0	80,0	41,0	45,0	30,0	28,0
Водопостачання; каналізація, обробка відходів	235,0	202,0	76,0	180,0	65,0	106,0	133,0	32,0	38,0	15,0	23,0
Будівництво	972,0	602,0	442,0	852,0	290,0	503,0	604,0	245,0	225,0	113,0	113,0
Оптова і рознична торгівля; ремонт автотранспортних засобів і мотоциклів	3119,0	2322,0	1546,0	2243,0	952,0	1704,0	1471,0	720,0	586,0	518,0	569,0
Транспорт, складське господарство, поштова і кур'єрська діяльність	612,0	490,0	316,0	546,0	176,0	300,0	391,0	149,0	129,0	83,0	105,0
Тимчасове розміщення і організація харчування	573,0	528,0	259,0	309,0	144,0	254,0	170,0	70,0	70,0	49,0	38,0
Інформація і телекомунікації	894,0	734,0	428,0	616,0	344,0	594,0	392,0	170,0	139,0	154,0	160,0
Операції з нерухомим майном	456,0	305,0	185,0	379,0	149,0	244,0	303,0	106,0	104,0	64,0	64,0
Професійна, наукова і технічна діяльність	787,0	545,0	399,0	623,0	277,0	470,0	408,0	220,0	159,0	149,0	154,0
Діяльність в сфері адміністративного і допоміжного обслуговування	670,0	506,0	314,0	528,0	210,0	337,0	380,0	164,0	145,0	107,0	107,0
Надання інших видів послуг	23,0	22,0	15,0	19,0	8,0	15,0	13,0	5,0	3,0	7,0	7,0
Max	3119,0	2322,0	1546,0	2243,0	952,0	1704,0	1471,0	720,0	586,0	518,0	569,0
Min	23,0	22,0	15,0	19,0	8,0	15,0	13,0	5,0	3,0	7,0	7,0

Частные показатели были стандартизированы с учетом типа воздействия каждого на развитие социальных медиа и больших данных; рассчитан вектор матрицы стандартизированных данных. Для расчета интегрального показателя оценки развития социальных медиа и больших данных (рис. 1) использован метод энтропии.

На третьем этапе проводится идентификация влияния каждого частного показателя на изменения интегрального показателя с помощью регрессионного анализа – построения уравнения множественной линейной регрессии.

Уравнение множественной линейной регрессии имеет вид:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n, \quad (1)$$

где α – свободный член, показывает влияние посторонних факторов, не включенных в

модель; β_1, β_n – коэффициенты, отражающие влияние независимых переменных на изменения зависимой переменной; y – зависимая переменная; x_1, x_n – независимые переменные.

В данном исследовании переменная, y (интегральный показатель оценки развития социальных медиа и больших данных в Украине, $I_{soc-bigdata}$) зависит от частных показателей, объясняющих его (i.1 – i.11), $n = 11$. Регрессионная модель является линейной относительно ее коэффициентов, β .

Коэффициенты модели, β , рассчитываются по методу наименьших квадратов, который минимизирует сумму квадратов прогнозных ошибок (обозначается также как Sum of Squared Errors), α – свободный член модели обозначается как сдвиг модели; его математическая интерпретация – определять прогнозное значение, y , при условии, что все независимые переменные = 0.

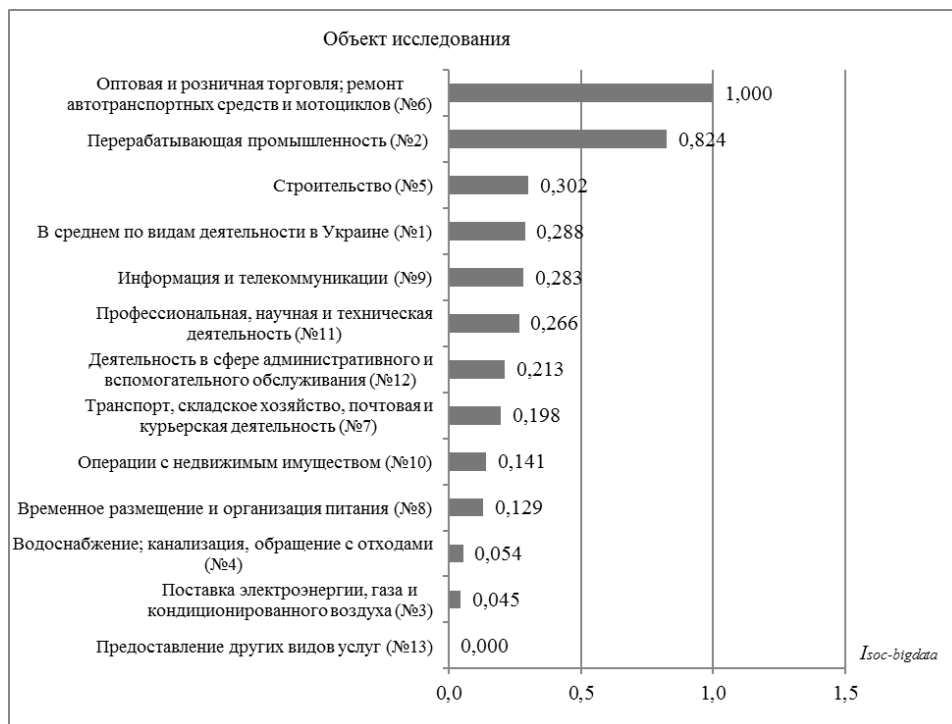


Рис. 1. Результат реализации первого и второго этапов метода дидактической редукции в MS Excel: распределение значений интегрального показателя оценки развития социальных медиа и больших данных в Украине ($I_{soc-bigdata}$) (далее в статье будут использованы только номера объектов исследования)

MS Excel позволяет рассчитать параметры модели, β , с помощью функции «=ЛИНЕЙН ()». Необходимо рассчитать параметры линейного приближения модели по методу наименьших квадратов, для этого следует реализовать следующие действия: выделить диапазон ячеек, где будут

размещены значения параметров (равен количеству независимых переменных, для данного исследования = 11) и в строке формул задать функцию следующего вида:

=ЛИНЕЙНЫЕ

(известные_значения_y;[известные_значения_x];ИСТИНА) (рис. 2).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		i.1	i.2	i.3	i.4	i.5	i.6	i.7	i.8	i.9	i.10	i.11		Isoc-bigdata
2	№1	0,2963	0,2897	0,2744	0,3072	0,2949	0,2932	0,3102	0,29	0,299	0,258	0,2527		0,287921
3	№2	0,8953	0,80783	0,7655	0,9033	0,8114	0,8786	0,8409	0,869	0,822	0,72	0,7349		0,823744
4	№3	0,0404	0,04435	0,0255	0,0382	0,0424	0,0485	0,046	0,05	0,072	0,045	0,0374		0,044567
5	№4	0,0685	0,07826	0,0398	0,0724	0,0604	0,0539	0,0823	0,038	0,06	0,016	0,0285		0,054434
6	№5	0,3065	0,25217	0,2789	0,3746	0,2987	0,2889	0,4053	0,336	0,381	0,207	0,1886		0,302257
7	№6	1	1	1	1	1	1	1	1	1	1	1		1
8	№7	0,1902	0,20348	0,1966	0,237	0,178	0,1687	0,2593	0,201	0,216	0,149	0,1744		0,197754
9	№8	0,1776	0,22	0,1594	0,1304	0,1441	0,1415	0,1077	0,091	0,115	0,082	0,0552		0,12947
10	№9	0,2813	0,30957	0,2698	0,2684	0,3559	0,3428	0,2599	0,231	0,233	0,288	0,2722		0,28283
11	№10	0,1399	0,12304	0,111	0,1619	0,1494	0,1356	0,1989	0,141	0,173	0,112	0,1014		0,140835
12	№11	0,2468	0,22739	0,2508	0,2716	0,285	0,2694	0,2709	0,301	0,268	0,278	0,2616		0,266458
13	№12	0,209	0,21043	0,1953	0,2289	0,214	0,1906	0,2517	0,222	0,244	0,196	0,1779		0,212756
14	№13	0	0	0	0	0	0	0	0	0	0	0		0
15														
16		0,0896	0,08856	0,0898	0,0936	0,0904	0,0935	0,0899	0,094	0,089	0,089	0,0934		
17														

Рис. 2. Расчет параметров линейного приближения модели по методу наименьших квадратов

Таким образом, в диапазоне В16:L16 получены параметры линейного приближения модели (не стандартизированные коэффициенты, β).

Однако для дальнейшего расчета коэффициентов, β , множественной регрессионной модели пошаговые расчеты значительно усложняются, так как необходимо применять матричный подход.

Следует также отметить, что значения параметров модели по такому способу

расчета отражаются в выделенном диапазоне ячеек в обратном порядке (в ячейке L16 отражено значение параметра при частном показателе i.1). Поэтому для большего удобства рекомендуется использовать дополнительную настройку MS Excel пакет «Анализ данных».

Активировав данный пакет, выбирают функцию «регрессия», задают параметры анализа и получают результаты (рис. 3).

	A	B	C	D	E	F	G	H	I	
1	Вывод ИТОВ									
2										
3	Регрессионная статистика									
4	Множественный R	1								
5	R-квадрат	1								
6	Нормированный R-квад	1								
7	Стандартная ошибка	7,48E-18								
8	Наблюдения	13								
9										
10	Дисперсионный анализ									
11		df	SS	MS	F	Значимость F				
12	Регрессия	11	1,052011	0,095637	1,71E+33	1,88681E-17				
13	Остаток	1	5,6E-35	5,6E-35						
14	Итого	12	1,052011							
15										
16		Коэффиц	Стандар	t-	P-	Верхние	Нижние	Верхние		
		иенты	тная	статист	Значение	Нижние 95%	95%	95,0%	95,0%	
			ошибка	ика						
17	Y-пересечение	4,16E-17	7,46E-18	5,583607	0,11282	-5,31086E-17	1,36E-16	-5,3E-17	1,36E-16	
18	Переменная X 1	0,093382	5,39E-16	1,73E+14	3,68E-15	0,093381583	0,093382	0,093382	0,093382	
19	Переменная X 2	0,088868	4,12E-16	2,16E+14	2,95E-15	0,088867677	0,088868	0,088868	0,088868	
20	Переменная X 3	0,088821	2,85E-16	3,11E+14	2,04E-15	0,088821364	0,088821	0,088821	0,088821	
21	Переменная X 4	0,093735	1,37E-15	6,82E+13	9,33E-15	0,093735392	0,093735	0,093735	0,093735	
22	Переменная X 5	0,089907	7,3E-16	1,23E+14	5,17E-15	0,089907301	0,089907	0,089907	0,089907	
23	Переменная X 6	0,0935	7,02E-16	1,33E+14	4,78E-15	0,093499916	0,0935	0,0935	0,0935	
24	Переменная X 7	0,090351	8,11E-16	1,11E+14	5,72E-15	0,090350773	0,090351	0,090351	0,090351	
25	Переменная X 8	0,093565	7,64E-16	1,22E+14	5,2E-15	0,093564575	0,093565	0,093565	0,093565	
26	Переменная X 9	0,089754	4,71E-16	1,91E+14	3,34E-15	0,089754247	0,089754	0,089754	0,089754	
27	Переменная X 10	0,088559	9,17E-16	9,66E+13	6,59E-15	0,088559357	0,088559	0,088559	0,088559	
28	Переменная X 11	0,089558	5,26E-16	1,7E+14	3,74E-15	0,089557815	0,089558	0,089558	0,089558	
29										

Рис. 3. Окно результатов регрессионного анализа в MS Excel

Таким образом, с помощью MS Excel возможно провести регрессионный анализ двумя способами, но при этом следует учитывать, что с помощью данного средства

можно получить значение «естественных» коэффициентов регрессии. Для более точных результатов необходимо проводить дополнительные расчеты стандартизованных

коэффициентов с учетом дисперсий значений независимых переменных ($i.1 - i.11$) и зависимой переменной y ($I_{soc-bigdata}$).

На четвертом этапе метода производится группировка объектов исследования с помощью кластерного анализа по частным показателям оценки развития социальных медиа и больших данных.

Процедура кластерного анализа в MS Excel достаточно трудоемкая. Проведем ее на примере кластеризации 13 объектов исследования по двум частными показателями оценки развития социальных медиа $i.1$ и $i.2$ (рис. 4). Кластеризация осуществлена на основе вычисления евклидового расстояния между объектами.

Рис. 4. Пошаговое заполнение матрицы евклидовых расстояний между объектами исследования

Заполнив матрицу евклидовых расстояний между объектами исследования (расстояние между объектами 1 и 2; 2 и 3; 3 и 4 и т.д.), можно вручную осуществить объединение в кластеры тех объектов, которые имеют наименьшее евклидовое расстояние.

Из рис. 4 видно, что наиболее близки друг к другу объекты №7 и №12

(минимальное евклидово расстояние в матрице = 0,019984). Целесообразным является объединение их в один кластер, оставляя наименьшее значение расстояния в соответствующих строке и столбце. Матрица уменьшается на одну строку и один столбец (рис. 5).

	№1	№2	№3	№4	№5	№6	№7, №12	№8	№9	№10	№11	№13
№1	0	0,792045	0,354522	0,310809	0,038894	0,999873	0,117918	0,137594	0,024869	0,228566	0,079584	0,414377
№2	0,792045	0	1,146246	1,102717	0,809607	0,218821	0,9099338	0,927703	0,790748	1,019654	0,870379	1,205916
№3	0,354522	1,146246	0	0,044043	0,337679	1,354309	0,218595	0,22293	0,35833	0,126846	0,275869	0,059974
№4	0,310809	1,102717	0,044043	0	0,29481	1,310474	0,1746635	0,17891	0,314339	0,084267	0,232441	0,103989
№5	0,038894	0,809607	0,337679	0,29481	0	1,019878	0,1061001	0,132831	0,062678	0,210837	0,06469	0,396924
№6	0,999873	0,218821	1,354309	1,310474	1,019878	0	1,1176435	1,133429	0,996587	1,228372	1,079018	1,414214
№7, №12	0,117918	0,909934	0,218595	0,174663	0,1061	1,135847	0	0,020776	0,122726	0,094914	0,041421	0,278562
№8	0,137594	0,927703	0,22293	0,17891	0,132831	1,133429	0,020776	0	0,137011	0,104061	0,069516	0,28277
№9	0,024869	0,790748	0,35833	0,314339	0,062678	0,996587	0,122726	0,137011	0	0,234105	0,089146	0,418303
№10	0,228566	1,019654	0,126846	0,084267	0,210837	1,228372	0,094914	0,104061	0,234105	0	0,149394	0,186279
№11	0,079584	0,870379	0,275869	0,232441	0,06469	1,079018	0,0414205	0,069516	0,089146	0,149394	0	0,335563
№13	0,414377	1,205916	0,059974	0,103989	0,396924	1,414214	0,2785619	0,28277	0,418303	0,186279	0,335563	0

Рис. 5. Первый шаг объединения объектов исследования в кластеры

В один кластер можно объединить объекты №7, №12 и №9. Проводя такие пошаговые операции, можно получить итоговый результат кластеризации объектов исследования.

Таким образом, в MS Excel был проведен кластерный анализ по двум частными показателями методом «ближайшего соседа» и получено два кластера, евклидово расстояние между которыми составляет 0,790748. Однако в исследовании

анализируется более двух частных показателей, что усложняет процедуру и не позволяет рекомендовать средство MS Excel, как наиболее оптимальное для использования в рассматриваемых целях процесса обучения.

3.2. Использование MS Excel и STATISTICA на этапах реализации метода

Согласно варианту 2, реализация первого и второго этапов метода проводится в программе Excel, а третьего и четвертого – в программе STATISTICA.

На третьем этапе для проведения регрессионного анализа в окне рабочей книги формируется матрица для анализа, где

строками являются объекты исследования (виды экономической деятельности, №1-№13), а столбцами – стандартизированные частные и интегральный показатель оценки развития социальных медиа и больших данных в Украине.

Для регрессионного анализа необходимо выбрать вкладку «Анализ → Углубленные методы → Множественная регрессия», и в выпадающем окне выбрать независимые (i.1 – i.11) и зависимую ($I_{soc-bigdata}$) переменные, метод расчета (backward stepwise) (рис. 6) и получить результаты анализа.

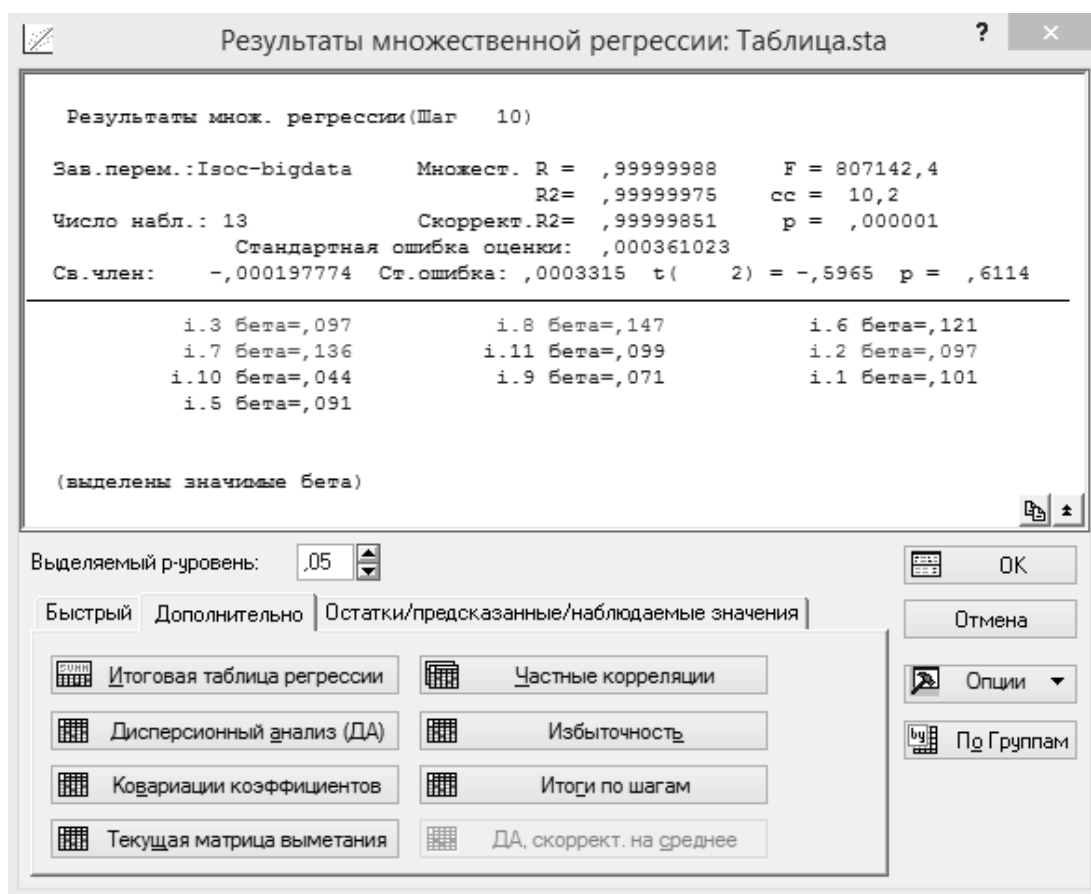


Рис. 6. Регрессионный анализ в STATISTICA

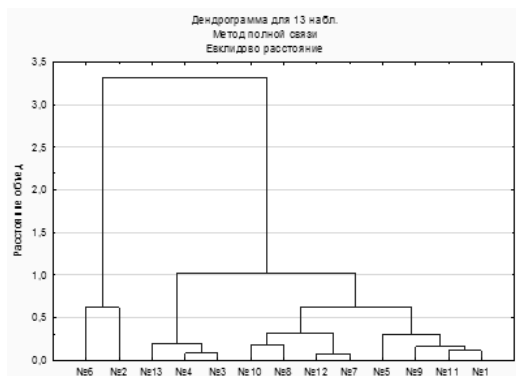
В результатах представлены стандартизированные коэффициенты модели. Недостатком в данном виде представления результатов является не всегда корректное выделение значимых бета модели. Например, если не выделены цветом бета при i.1 и i.6, следовательно, выводы требуют дополнительного внимания.

Данная программа позволяет получать график распределения остатков регрессионной модели, описательную статистику (средние, стандартные отклонения, корреляционную матрицу), а также прогнозировать изменения

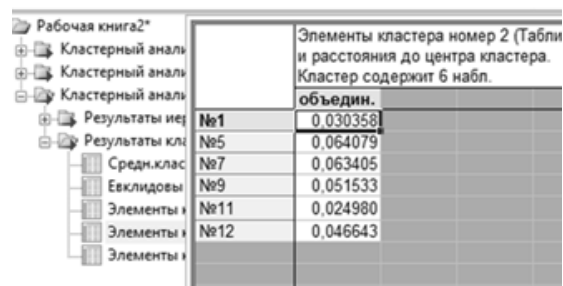
зависимой переменной в будущем.

На четвертом этапе для проведения кластерного анализа необходимо выбрать вкладку «Анализ → Многомерный анализ → Кластерный анализ». Предлагается проведение трех методов кластеризации: иерархическая классификация, k-средних и двухходового объединения.

Результаты кластеризации иерархическим методом (кластерных деревьев) по правилу полной связи позволяют получить дендрограмму объединения объектов в кластеры (рис. 7).



а. Методом кластерных деревьев



б. Методом k-средних

Рис. 7. Результаты кластерного анализа в STATISTICA

Пользователю необходимо только самостоятельно определить расстояние, на котором будет осуществлено объединение объектов. Так, на расстоянии 0,99 можно выделить три кластера, а на расстоянии 1,50 – два кластера. Еще более простым в интерпретации является метод k-средних, который позволяет задавать количество кластеров самостоятельно и получать перечень объектов, попавших в каждый кластер.

Таким образом, по характеру полученных результатов, скорости и производительности вариант 2 использования средств программного обеспечения превосходит вариант 1.

3.3. Использование MS Excel и SPSS на этапах реализации метода

Согласно варианту 3, реализация первого и второго этапов метода проводится в программе Excel, а третьего и четвертого – в программе SPSS.

В исходном окне анализа была введена совокупность исследуемых показателей и объектов. Далее следует выбрать команды «Анализ → Регрессии → Линейная» и выбрать зависимые и независимые переменные, а также метод (пошагово назад) (рис. 8).

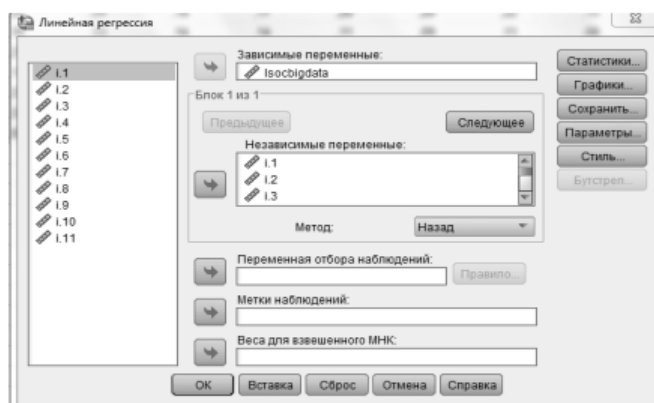


Рис. 8. Окно задания параметров регрессионной модели в SPSS

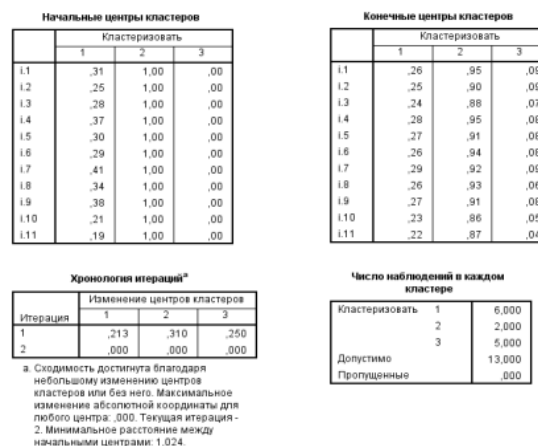
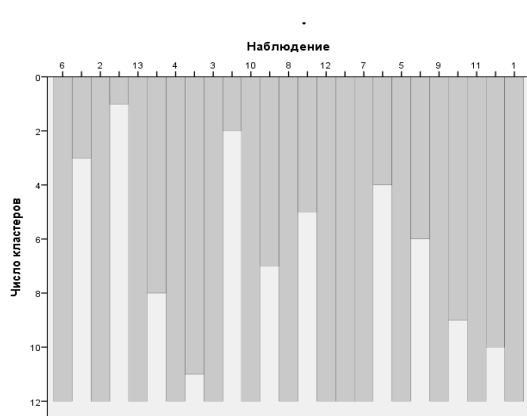
Результаты анализа представляют описательные статистики для модели, матрицу корреляций, параметры модели, стандартизованные и не стандартизованные коэффициенты бета в табличном виде.

Визуальное представление результатов регрессионного анализа в STATISTICA более наглядное и легкое в интерпретации, чем в SPSS, поскольку SPSS выдает результаты в виде полотна выводов, а STATISTICA – в виде окна с наиболее значимыми результатами и

кнопками, по которым можно перейти к более детальным описаниям по принципу матрешки.

На четвертом этапе метода проводится кластеризация объектов исследования.

Программа SPSS предлагает три метода кластеризации: двустороннего присоединения («двухэтапный кластерный»), иерархическая классификация (кластерных деревьев) и метод k-средних. На рис. 9 представлены результаты кластеризации объектов исследования методом иерархической классификации и k-средних.



а. Методом иерархической классификации

б. Методом k-средних

Рис. 9. Результаты кластерного анализа в SPSS

Рис. 9 показывает, что представление результатов кластерного анализа в SPSS имеет несколько худший вид, их труднее интерпретировать по сравнению с формой представления результатов в STATISTICA. Например, представление результатов иерархической кластеризации в виде дендрограммы (STATISTICA), является более наглядным, чем в виде диаграммы (SPSS).

3.4. Результаты дидактической редукции содержания обучения развития социальных медиа и больших данных.

Таким образом, применение метода дидактической редукции в педагогической практике позволило распределить по значимости изучение развития социальных медиа и больших данных на предприятиях, которые специализировались на разных видах экономической деятельности в Украине в 2018 году с помощью интегрального показателя, и выявить, что наибольшее влияние на изменение интегрального показателя оказывали такие частные:

- представление предприятия или рекламирование его работы (товаров, услуг) в социальных медиа (i.1), $\beta = 0,101$;
- получение отзывов клиентов или предоставление ответов на их вопросы в социальных медиа (i.2), $\beta = 0,097$;
- привлечение клиентов в развитие или инновацию товаров и услуг в социальных медиа (i.3), $\beta = 0,097$;
- блоги или микроблоги предприятия в социальных медиа (i.5), $\beta = 0,091$;
- вебсайты с мультимедийным содержанием в социальных медиа (i.6), $\beta = 0,121$;
- средства обмена знаниями в социальных медиа (i.7), $\beta = 0,136$;
- развитие больших данных своего предприятия, полученных со смарт-устройств

или датчиков геолокационных данных (i.8), $\beta = 0,147$;

- развитие больших данных, полученных с портативных устройств (i.9), $\beta = 0,071$;
- развитие больших данных, сформированных из социальных медиа (i.10), $\beta = 0,044$.

Другие показатели, сотрудничество с деловыми партнерами или другими организациями в социальных медиа (i.4) и развитие больших данных, сформированных из других источников (i.11), не оказывают статистически значимого влияния на интегральный показатель развития социальных медиа и больших данных в Украине и могут анализироваться в процессе обучения после показателей, оказывающих большее влияние, в произвольном порядке при условии наличия доступного учебного времени.

Кластерный анализ привел к объединению объектов исследования в кластеры и определения тех видов деятельности, которые являются характерными представителями каждого кластера.

В кластер 1, который характеризуется высоким уровнем развития социальных медиа и больших данных, попали такие виды деятельности, как перерабатывающая промышленность и оптовая и розничная торговля; ремонт автотранспортных средств и мотоциклов. Оба объекта имеют одинаковое расстояние до центра кластера (0,093606).

В кластер 2, характеризующийся средним уровнем развития социальных медиа и больших данных, попали 6 объектов исследования, в том числе предприятия строительства, транспорта, складского хозяйства, почтовой и курьерской деятельности и другие. Характерным представителем кластера является такой вид

деятельности, как профессиональная, научная и техническая деятельность.

В кластер 3, характеризующийся низким уровнем развития социальных медиа и больших данных, попали 5 объектов исследования, в том числе поставка электроэнергии, газа и кондиционированного воздуха; водоснабжение; канализация, обращение с отходами; временное размещение и организация питания. Характерным представителем кластера является такой вид деятельности, как водоснабжение; канализация, обращение с отходами.

В процессе учебной деятельности целесообразно анализировать всех представителей кластера последовательно, начиная с характерного представителя кластера. Кроме того, формирование трех кластеров дает возможность оценивать частные показатели оценки развития социальных медиа и больших данных в системе «плохо – нормально – хорошо», что соответствует низкому (кластер 3), среднему (кластер 2) и высокому (кластер 3) уровням.

4. Выводы и перспективы дальнейших исследований. Реализация предложенного метода дидактической редукции тремя

вариантами применения средств программного обеспечения позволила оценить их по предложенным критериям. Наиболее эффективным оказался второй вариант применения средств программного обеспечения – использование MS Excel на первом и втором, а STATISTICA на третьем и четвертом этапах проведения метода. По некоторым критериям ему уступает вариант использования MS Excel на первом и втором, а SPSS на третьем и четвертом этапах реализации метода дидактической редукции.

В целом приведенный пример редукции содержания обучения развития социальных медиа и больших данных свидетельствует о целесообразности применения метода при формировании программ курсов и отдельных занятий с использованием рассмотренных средств ИКТ.

Перспектива дальнейших исследований состоит в подтверждении данных результатов с привлечением статистически значимой выборки студентов, а также в изучении возможностей реализации метода дидактической редукции с помощью программных продуктов, ориентированных на программирование, например, R и MATLAB.

Список использованной литературы:

1. Syamsuri B. S. Development of Teaching Material Oxidation-Reduction Reactions through Four Steps Teaching Material Development (4S TMD) / B. S. Syamsuri, S. Anwar, O. Sumarna // *Journal of Physics: Conference Series*. – 2017. – Vol. 895, No. 1. – Pp. 15-24.
2. Lehner M. Didaktische Reduktion / M. Lehner. – Bern : Haupt, 2012. – 211 s.
3. Shmatkov D. The use of causal maps as interdisciplinary didactic reduction method / D. Shmatkov // *Education*. – 2016. – No. 6. – Pp. 16-21.
4. The IMPACT2C web-atlas – Conception, organization and aim of a web-based climate service product / S. Preuschmann [et al.] // *Climate Services*. – 2017. – No. 7. – Pp. 115–125.
5. Costa G. GraphsJ 3: A modern didactic application for graph algorithms / G. Costa, C. D'Ambrosio, S. Martello // *Journal of Computer Science*. – 2014. – No. 10 (7). – Pp. 1115–1119.
6. Krpan D. Posredovani prijenos u poučavanju programiranja s vizualnim programskim jezicima / D. Krpan, D. Brčić // *Politehnika : Časopis za tehnički odgoj i obrazovanje*. – 2018. – No. 2 (1). – Pp. 71–79.
7. Hazzan O. Guide to teaching computer science: An activity based approach / O. Hazzan, T. Lapidot, N. Ragonis. – New-York : Springer. – 345 p.
8. Ressourcenökonomische Erstellung von Materialien für Lehrende und Lernende in der Studiengangsphase / T. Bentz, A. Helfrich-

Schkarbanenko, R. Кош [und andere]// HD MINT Symposium. – Nürnberg, 2015. – S. 87–92.

9. Developing an environmental monitoring program based on the principles of didactic reduction / D. Shmatkov, N. Bielikova, N. Antonenko [et. al.] // *European Journal of Geography*. – 2019. – No. 10 (1). – Pp. 99–116.

10. Шматков Д. Редукція змісту навчання моніторингу стану ґрунту середовища існування / Д. Шматков, Н. Белікова, О. Шелковий // *Молодь і ринок*. – 2019. – Вип. 2 (169). – С. 13–19.

11. Шматков Д. І. Формування змісту навчання моніторингу атмосферного повітря середовища існування на засадах дидактичної редукції / Д. І. Шматков, О. О. Шелковий // *Педагогіка безпеки*. – 2018. – Вип. 3 (2). – С. 160–170.

12. Easy methods for extracting individual regression slopes: Comparing SPSS, R, and Excel / R. Pfister, K. Schwarz, R. Carson [at el.] // *Tutorials in Quantitative Methods for Psychology*. – 2013. – No. 9 (2). – Pp. 72–78.

13. Abbott M. L. Using statistics in the social and health sciences with SPSS and excel / M. L. Abbott. –: John Wiley & Sons, 2016. – 527 p.

14. Державна служба статистики, Використання інформаційно-комунікаційних технологій на підприємствах у 2018 році. [Електронний ресурс]. – Режим доступу : <http://www.ukrstat.gov.ua/> - Дата звернення 01.03.2020р.

References:

1. Syamsuri, BS, Anwar, S & Sumarna, O 2017, 'Development of Teaching Material Oxidation-Reduction Reactions through Four Steps Teaching Material Development (4S TMD)', Journal of Physics: Conference Series, vol. 895, no. 1, pp. 15-24.
2. Lehner, M 2012, Didaktische Reduktion, Haupt, Bern.
3. Shmatkov, D 2016, 'The use of causal maps as interdisciplinary didactic reduction method', Education, no. 6, 16-21.
4. Preuschmann, S et al. 2017, 'The IMPACT2C web-atlas – Conception, organization and aim of a web-based climate service product', Climate Services, no. 7, pp. 115-125.
5. Costa, G, D'Ambrosio, C & Martello, S 2014, 'GraphsJ 3: A modern didactic application for graph algorithms', Journal of Computer Science, no. 10 (7), pp. 1115-1119.
6. Krpan, D & Brčić, D 2018, 'Posredovani prijenos u poučavanju programiranja s vizualnim programskim jezicima', Politehnika : Časopis za tehnički odgoj i obrazovanje, no. 2 (1), pp. 71-79.
7. Hazzan, O, Lapidot, T & Ragonis, N 2015, Guide to teaching computer science: An activity based approach, Springer, New-York.
8. Bentz, T, Helfrich-Schkarbanenko, A, Koß, R & Nitsche, A 2015, 'Ressourcenökonomische Erstellung von Materialien für Lehrende und Lernende in der Studieneingangsphase', HD MINT Symposium, Technische Hochschule, Nürnberg, pp. 87-92.
9. Shmatkov, D, Bielikova, N, Antonenko, N & Shelkovyj, A 2019, 'Developing an environmental monitoring program based on the principles of didactic reduction', European Journal of Geography, no. 10 (1), pp. 99-116.
10. Shmatkov, D, Bielikova, N & Shelkovyj, A 2019, 'Reduktsiia zmistu navchannia monitorynhu stanu gruntu seredovysheha isnuvannia'[Reducing the content of training to monitor the condition of the soil habitat], Molod i rynek, iss. 2 (169), pp. 13-19.
11. Shmatkov, D & Shelkovyj, A 2018, 'Formuvannia zmistu navchannia monitorynhu atmosferneho povitria seredovysheha isnuvannia na zasadakh dydaktychnoi reduksii'[Formation of the content of the teaching of atmospheric air monitoring of the living environment based on didactic reduction], Pedahohika bezpeky, iss. 3 (2), pp. 160-170.
12. Pfister, R, Schwarz, K, Carson, R & Janczyk, M 2013, 'Easy methods for extracting individual regression slopes: Comparing SPSS, R, and Excel', Tutorials in Quantitative Methods for Psychology, no. 9 (2), pp. 72-78.
13. Abbott, ML 2016, Using statistics in the social and health sciences with SPSS and excel. John Wiley & Sons, город.
14. Derzhavna sluzhba statystyky, Vykorystannia informatsiino-komunikatsiinykh tekhnolohii na pidprijemstvakh u 2018 rotsi 2018, [State Statistics Service, Utilization of information and communication technologies in enterprises in 2018] viewed 01 March 2020, <<http://www.ukrstat.gov.ua/>>.

Статья поступила в редакцию 15.03.2020 г.