

Khassenova Maral Torgayevna – PhD, Senior Lecturer of the “Chemistry” educational program, Alkey Margulan Pavlodar Pedagogical University, Republic of Kazakhstan, 140000 Pavlodar, 60 Olzhabai batyr Str., tel.: +7-777-975-45-06, e-mail: maral_176@mail.ru.*

[Khamzina Sholpan Shapiyevna] – Candidate of Pedagogical Sciences, Professor of the “Chemistry” educational program, Alkey Margulan Pavlodar Pedagogical University, Republic of Kazakhstan, 140000 Pavlodar, 60 Olzhabai batyr Str., e-mail: khamzina_64@mail.ru.

XҒТАР 14.85.09

ӨОЖ 378.147.004.04

<https://doi.org/10.52269/RWEP2522187>

ЖОҒАРЫ ОҚУ ОРНЫНДА БІЛІМ АЛУШЫЛАРДЫ ҮЛКЕН КӨЛЕМДІ ДЕРЕКТЕР БОЙЫНША ДАЯРЛАУДЫҢ ПРАКТИКАЛЫҚ НЕГІЗДЕРІ (R ПРОГРАММАЛАУ ТІЛІНДЕ ҮЛКЕН КӨЛЕМДІ ДЕРЕКТЕРДІ БӨЛІКТЕРМЕН ӨНДЕУ)

Нурбекова Г.Ф. – PhD, информатика кафедрасының қауымдастырылған профессор м.а., Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан Республикасы.

Ерланова Г.Ж. – PhD, ақпараттық-техникалық ғылымдары кафедрасының қауымдастырылған профессор м.а., Alikhan Bokeikhan university, Семей қ., Қазақстан Республикасы.

Зулпыхар Ж.Е. – педагогика ғылымдарының кандидаты, қауымдастырылған профессор, информатика кафедрасының меңгерушісі, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан Республикасы.*

Нариман С.А. – PhD, информатика кафедрасының қауымдастырылған профессор м.а., Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан Республикасы.

Мақалада үлкен деректер саласындағы мамандарды даярлау, R бағдарламалау ортасындағы үлкен деректерді өңдеу, сақтау және талдау негізінде жоғары оқу орындарындағы білім алушылардың білімін арттыру қарастырылады. Ұсынылған нәтижелер аппараттық және бағдарламалық қамтамасыз ету және бағдарламалау тілдері туралы білімді жан-жақты зерттеуге және білім беру мазмұнына енгізуге бағытталған зерттеу жұмысының бөлігі болып табылады. Студенттерді R тілінің пакеттерімен және деректерді сақтау форматтарымен таныстыруға ерекше назар аударылады. Деректерді екі түрлі жолмен .rds және .csv форматтарында сақтауға болатындығы және әрқайсысының үлкен деректерді өңдеу кезінде өзіндік ерекшеліктері мен артықшылықтары бар екендігі көрсетілген. Үлкен деректер құрылымдалған, жартылай құрылымдалған (XML және JSON) және құрылымдалмаған (мәтіндер, суреттер мен бейнелер) болып бөлінеді, бұл оларды сақтауды, өңдеуді және талдауды күрделендіреді. Міндет: R жадына толық деректер жиынтығын бірден жүктеу мүмкін болмаған жағдайын қарастыру. Үлкен деректерді талдау кезінде R жадына толық деректер жиынтығын бірден жүктеу мүмкін болмаған жағдайда деректерді фрагменттермен өңдеу мүмкіндігі және осы жағдайда Саймон Урбанек пен Тейлор Арнольдтың iotools пакетінің chunk.apply функциясының қолданылу жолдары қарастырылады. Үлкен көлемді деректер бойынша жоғары оқу орнында білім алушыларды даярлауға байланысты талдаулар жасалып, біздің зерттеу жұмысымыздың практикалық түрде жүзеге асырылуының нәтижелік жұмыстарынан мәліметтер келтірілді.

Түйінді сөздер: үлкен көлемді деректер, үлкен көлемді деректер туралы талдаулар, R программалау ортасы, Nycflights13 пакеті, flights файлдары.

ПРАКТИЧЕСКИЕ ОСНОВЫ ПОДГОТОВКИ ОБУЧАЮЩИХСЯ В ВУЗЕ ПО БОЛЬШИМ ДАННЫМ (ОБРАБОТКА БОЛЬШИХ ДАННЫХ ЧАСТЯМИ НА ЯЗЫКЕ ПРОГРАММИРОВАНИЯ R)

Нурбекова Г.Ф. – PhD, и.о. ассоциированного профессора кафедры информатики, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Республика Казахстан.

Ерланова Г.Ж. – PhD, и.о. ассоциированного профессора кафедры информационно-технических наук, Alikhan Bokeikhan university, г. Семей, Республика Казахстан.

Зулпыхар Ж.Е. – кандидат педагогических наук, ассоциированный профессор, заведующий кафедрой информатики, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Республика Казахстан.*

Нариман С.А. – PhD, и.о. ассоциированного профессора кафедры информатики, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Республика Казахстан.

В статье рассматривается подготовка специалистов в области больших данных, повышение знаний обучающихся в высших учебных заведениях на основе обработки, хранения и анализа больших данных в среде программирования R. Представленные результаты являются частью

исследовательской работы, направленной на всестороннее изучение и внедрение знаний в содержание материалов об аппаратно-программном обеспечении и языках программирования. Особое внимание уделяется ознакомлению студентов с пакетами языка R и форматами хранения данных. Показано, что данные могут быть сохранены двумя разными способами – в форматах `.rds` и `.csv`, каждый из которых имеет свои особенности и преимущества при последующей обработке больших данных. Большие данные делятся на структурированные, полуструктурированные (XML и JSON) и неструктурированные (тексты, изображения и видео), что затрудняет их хранение, обработку и анализ. Задача: рассмотреть случай, когда невозможно сразу загрузить полный набор данных в память R, возможность обработки данных фрагментами при невозможности сразу загрузить полный набор данных в память R при анализе больших данных, в данном случае упоминается использование функции `chunk.apply` из пакета `iotools` Саймона Урбанека и Тейлора Арнольда. Проведен анализ по большим данным, связанный с подготовкой обучающихся в вузе, приведены данные результативной работы из практической части нашей исследовательской работы.

Ключевые слова: большие данные, анализ больших данных, язык программирования R, пакет `Nycflights13`, файлы `flights`.

PRACTICAL BASICS OF TRAINING UNDERGRADUATES IN BIG DATA (PROCESSING BIG DATA IN CHUNKS USING THE R PROGRAMMING LANGUAGE)

Nurbekova G.F. – PhD, acting Associate Professor of the Department of computer science, L.N. Gumilyov Eurasian National University, Astana, Republic of Kazakhstan.

Yerlanova G.Zh. – PhD, acting Associate Professor of the Department of information and technical sciences, Alikhan Bokeikhan university, Semey, Republic of Kazakhstan.

Zulpykhar Zh.Y.* – Candidate of Pedagogical Sciences, Associate Professor, Head of the Department of computer science, L.N. Gumilyov Eurasian National University, Astana, Republic of Kazakhstan.

Nariman S.A. – PhD, acting Associate Professor of the Department of computer science, L.N. Gumilyov Eurasian National University, Astana, Republic of Kazakhstan.

The article considers the training of specialists in big data, increasing the knowledge of students in higher educational institutions based on the processing, storage and analysis of big data in the R programming environment. The presented results are part of a research project aimed at a comprehensive study and integration of knowledge related to hardware-software systems and programming languages. Special attention is given to familiarizing students with R language packages and data storage formats. It is demonstrated that data can be stored in two different formats—`.rds` and `.csv`—each offering distinct features and advantages for subsequent big data processing. Big data is divided into structured, semi-structured (XML and JSON) and unstructured (texts, images, and videos), which makes their storage, processing, and analysis more complex. Objective: to consider the case when it is impossible to immediately load a complete set of data into R memory. The ability to process data in fragments when it is impossible to immediately load the full data set into R memory when analyzing big data, and in this case the use of the `chunk.apply` function from the `iotools` package by Simon Urbanek and Taylor Arnold is mentioned. The analysis of big data related to the training of undergraduates is carried out, data from the results of the practical part of our research is presented.

Key words: big data, big data analysis, R programming environment, `Nycflights13` package, `flights` files.

Қысқашы. Қазіргі таңда кез-келген дамыған елдерде сапалы білім беру жүйесі жүйелі түрде қалыптасқан. Жалпы, білім беру жүйесінің негізгі мақсаты – ғылым және білім жетістіктері негізінде жеке адамды кәсіби шыңдауға және қалыптастыруға арналған тиісті жағдайлар жасау, білім беруді ақпараттандыру, оқытудың жаңа технологиясын енгізу, халықаралық ғаламдық коммуникациялық желілерге шығу болып табылады. Біздің елімізде білім берудің барлық салаларында сапалы білім беру жүйесін қалыптастыру үшін, заман талабына сай оқу-әдістемелік бағдарламалар ұсыну маңызды болып табылады. Қазақстан Республикасының 2007 жылы 27 шілдеде №317 бұйрығына сай енгізілген «Білім туралы» Заңға сәйкес, «әр білім алушының жеке қабілетіне қарай дамуы, оның қабілетін, дарындылығын дамыту» сияқты маңызды мәселелер қарастырылады. Ал, мұндай мәселелерді дамыту үшін ақпараттық-технологияларды жетік меңгерген, кәсіби шыңдалған, білім берудің жүйесінің модернизациясына бейімделген жоғары білімді мамандар қажет [1].

Біздің зерттеу жұмысымыздың мақсаты болып отырған жоғары оқу орнында ақпараттық-коммуникациялық технологияларға байланысты болашақ педагог білім алушыларды үлкен көлемді деректер бойынша даярлаудың теориялық негіздерін анықтап және практикалық түрде жүзеге асыру, заман талабына сай, бәсекеге қабілетті болашақ маманды даярлаудың негізі болып табылады.

«Білімді ұлт» сапалы білім беру» ұлттық жобасында заманауи бәсекеге қабілетті мамандарды даярлаудың алғы шарттары қарастырылған [2]. Сонымен қатар, Қазақстан Республикасы Білім және

ғылым министрінің білім беру ұйымдары қызметінің үлгілік қағидаларында, жоғары және жоғары оқу орнынан кейінгі білімі бар кадрларды даярлау жөнінде бәсеке қабілетті мамандарды даярлау талаптары келтірілген [3].

R бағдарламалау тілінде үлкен көлемді деректермен жұмыс істеуге арналған пакеттермен танысып, үлкен деректерді талдау кезінде R жадына толық деректер жиынтығын бірден жүктеу мүмкін болмауы мүмкін. Бұл жағдайда деректерді фрагменттермен өңдеу мүмкіндігі өте пайдалы.

Үлкен деректер құрылымдалған, жартылай құрылымдалған және құрылымдалмаған болуы мүмкін. Құрылымдық деректер-бұл дәстүрлі кестелер мен мәліметтер базасы. Жартылай құрылымдалған файлдар, мысалы, XML немесе JSON файлдары. Құрылымдалмаған деректерге мамандандырылған талдау алгоритмдерін қажет ететін мәтіндер, суреттер, бейнелер кіреді.

Білім берудегі үлкен деректердің қайнар көздері: электрондық үлгерім журналдары, онлайн-тест нәтижелері, оқыту платформаларынан алынған мәліметтер (LMS), бейне дәрістер, студенттердің белсенділігінің сандық іздері.

Ғылыми және қолданбалы әдебиеттердегі "үлкен деректер" (Big Data) термині әдетте "5V" деп аталатын модель арқылы сипатталады: көлем (көлем), әртүрлілік (өзгергіштік), жылдамдық (жылдамдық), сенімділік (сенімділік) және мән (мән). Көлем дегеніміз-деректердің үлкен массивтері, әртүрлілік – олардың әртүрлі форматтары (мәтін, кескіндер, аудио, видео және т.б.), жылдамдық – қабылдау мен өңдеудің жеделдігі, сенімділік – деректердің дәлдігі мен сенімділігі, ал құндылық – шешім қабылдау үшін ақпараттың пайдалылығы. Осылайша, үлкен деректер тек ақпарат массиві ғана емес, сонымен қатар белгілі бір сақтау, талдау және өңдеу әдістерін қажет ететін ресурстар болып табылады.

Алдымен, үлкен көлемді деректер бойынша жоғары оқу орнында білім алушыларды даярлау мәселесіне тоқталсақ.

Білім беру саласындағы үлкен деректерге байланысты ойларды саралауда мынадай тұжырым жасауға болады: «Педагогтар үлкен көлемді деректерді қолдану барысында оқушылардың үлгерім жағдайын анықтайды, олардың жеткен жетістіктерін бағалайды, сонымен бірге білім беру жүйесіндегі қолданылатын бағалау мен қолдау сияқты жетілдірілген жүйені қолданады», – деп негіздейді [4, 76-84 б.].

Бүгінгі күні мамандар даярлауда білім мазмұнын ғылыми, техникалық, технологиялық, ақпараттық өзгерістерге қарай қадағалап, толықтырып отыру қажеттілігі туындайды.

Жоғары оқу орнының білім алушыларын даярлауда заманауи ақпараттық технологиялар кеңінен қолданылуда. Заманауи ақпараттық технологияларды оқытуда қолдану арқылы жоспарланған нәтижеге жету жолында IT инфрақұрылымының толығымен қауіпсіз және өнімді қызмет етуі, негізгі шарттардың бірі болып табылады. Өңделетін ақпарат көлемінің үнемі артып отыруы, өнімділік пен қауіпсіздікті арттыруға қойылатын талаптар да жоғарылай түсуде. Бұл талаптарды қанағаттандырудың ең тиімді тәсілі, оқытуға арналған ақпараттық технологияларды дамыту болып табылады [5, 8-16 б.].

Үлкен деректерді өңдеу әдістері мен технологияларының ішінде мыналар маңызды рөл атқарады:

- Hadoop-таратылған деректерді сақтау және өңдеу платформасы;
- MapReduce – әртүрлі серверлерде деректерді параллель өңдеуге мүмкіндік беретін бағдарламалау моделі;
- Spark-жедел жадта жұмыс істейтін MapReduce-ке жылдам балама;
- NoSQL-гетерогенді және тез жаңартылатын деректерді сақтауға арналған мәліметтер базасы (MongoDB, Cassandra және т.б.).

Бұл технологияларды пайдалану нақты уақыт режимінде үлкен көлемдегі ақпаратты талдау мәселелерін тиімді шешуге мүмкіндік береді.

1 кесте – Деректерді өңдеу әдістерін салыстырмалы талдау

Көрсеткіштер	Дәстүрлі деректерді өңдеу әдістері	Big Data технологиялары
Деректер көлемі	Кішкентай және орташа	Өте үлкен (терабайттар мен петабайттар)
Өңдеу жылдамдығы	Төмен немесе орташа	Жоғары, нақты уақытта өңдеу мүмкіндігі
Деректер түрлері	Құрылымдалған	Құрылымдалған, құрылымдалмаған, жартылай құрылымдалған
Икемділік	Төмен	Жоғары (динамикалық масштабталу)
Талдау мүмкіндігі	Шектеулі	Кең (жасанды интеллект, машиналық оқыту)
Қолдану саласы	Стандартты есептер	Болжамдық аналитика, білім беру аналитикасы, жеке траекториялар және т.б.

Үлкен деректерді өңдеу әдістері динамикалық жаңартылатын және құрылымы әртүрлі деректер талданатын білім беру тапсырмалары контекстінде жоғары бейімделуді, икемділікті, масштабталуды және тұрақтылықты қамтамасыз етеді.

Абдыкаримова А.Т. «Big Data: проблемы и технологии» атты мақаласында І.Жансүгіров атындағы Жетісу мемлекеттік университетінде [6, 223-226 б.] үлкен деректерге оқытуда блокчейн, бұлттық технологияларды қолдану сұрақтарына тоқталса, Кондратюк А. блокчейн мен үлкен деректердің маңызы туралы [7], Baalamurugan K., Prabu P., Vasanth N. және т.б. блокчейнді қолдайтын платформалар туралы [8, 1004-1-1004-13 б.] өз еңбектерінде келтіреді.

Біздің ұсынып отырған зерттеу жұмысымызда да аталған мәселелер қарастырылып өтеді, атап айтсақ, блокчейн технологиялары, бұлттық ресурстарды үлкен деректерді өңдеуде қолданудың практикалық негіздері, мысалы MapReduce технологиясын қолдануда жүзеге асырылады.

Деректерді талдауды оқу жоспарын бағалауға да қолдануға, талаптарға сәйкес келетін оқу бағдарламасын жасауда үлкен көлемді деректердің бұл мәселені шеше алатынын атап өтеді [9, 231 б.]. Сол сияқты, Bandy I.R., Zaman M. [10, 589-595 б.] академиялық ортадағы үлкен деректер: білім алушылардың үлгерімін арттыру үшін ұсынылатын негіз туралы айтады.

Моррисон А. үлкен көлемді деректер әкімшілердің, оқытушылар мен білім алушылардың өзара әрекеттесуін өзгертетінін, жоғарғы оқу орындарының әлеуетті білім алушыларды қалай қызықтыратынына, сондай-ақ оқытушылардың қазіргі білім алушыларға қалай қызмет ететініне әсер ететінін, сонымен қатар, жаңа технологиялар мен инновациялар болашақта білім беруде үлкен көлемді деректерді пайдаланудың жаңа мүмкіндіктерін жасайтынын атап өтеді [11, 22-29 б.].

Тесленко И.Б. «Big Data» атты еңбегінде Владимир мемлекеттік университетінде үлкен көлемді деректерді талдау арқылы оқытушылар білім алушылардың оқу кезінде қиын немесе сәтті болатын бағыттарын анықтай алады, білім алушылардың жеке оқу кезіндегі қажеттіліктерін түсінеді және жеке оқыту стратегиясын жасай алады. Бұл білім алушыларға білім беру траекторияларын таңдауға мүмкіндік береді. Мысалы, үлкен деректерді талдау күндізгі оқытудың дәстүрлі әдістері білім алушының үлгеріміндегі проблемаларға әкелетінін көрсетуі мүмкін. Сондай-ақ, деректер білім алушының онлайн режимінде алған білімінің деңгейін көрсетуі мүмкін. Бұл жағдайда профессор немесе кеңесші білім алушымен өздерінің ерекше оқу стиліне сәйкес келетін бағдарламаны немесе курсты таңдаудай алады» [12, 123 б.], – деп келтіреді.

Жоғары білім беру жүйесінде үлкен көлемді деректерді кеңінен қолдану білім алушыларды, әсіресе қауіпті адамдарға оқуға түсуге жол бермеуге көмектеседі. Мысалы, АҚШ пен Ұлыбританиядағы бірнеше университеттер мұқтаж білім алушыларды қолдау үшін алдын-ала ескерту жүйесі сияқты нәрсені жасау үшін білім алушылардың деректерін талдау жұмыстарын қолданады.

Чикаго университетінің аналитика саласындағы ғылым магистрі бағдарламасын Чикаго университетінің Грэм мектебі ұсынады. Білім алушылар үлкен деректерді талдау және бизнес-идеяларды құру қабілетін дамытады [13].

Жоғарыда келтірілген еңбектерді саралай отырып, үлкен деректерді әр салада қолдану болашақ құзыреттілігі жоғары мамандарды даярлауды талап ететінін көрсетеді.

Қытай Халық республикасының жоғары оқу орындарында үлкен деректерді оқыту туралы еңбекте Институт әзірлеген және кеңінен қолданылатын негізгі ақпараттық нұсқаулықтың компьютерлік платформасы журналының деректеріне негізделген, білім алушылардың кіруіне және платформадағы ресурстарды қарауға қатысты деректерді жинауы және алдын ала өңдеуі; одан кейін студенттердің платформаға кіруі және ресурстарды қарау жағдайына статистикалық талдау жасауында машиналық оқыту алгоритмі шешім ағаштарының әдістерін қолдануы факторлары қарастырылады. Талдау нәтижелеріне сәйкес оқытушылар оқу мазмұнын ұйымдастыра алады және оқушылардың оқу жағдайына негізделген оқу режимін құра алады [14, 309-313 б.].

Остиндегі Техас университетінде бизнес-аналитика саласындағы ғылым магистрі бағдарламасында білім алушыларға қаржы, маркетинг және жеткізілім тізбегін басқару сияқты бағыттар туралы түсінік алуға көмектесетін статистикалық талдау, деректерді іздеу, табиғи тілдерді өңдеу және машиналық оқыту дағдыларын ұсынады. Сондай-ақ, білім алушылар Walmart, Deloitte consulting, McKinsey салаларының көшбасшылары болып табылатын түлектермен қарым-қатынас жасау үшін керемет мүмкіндіктерді пайдалана алады [15].

Астана қаласының Astana IT University университетінің 6B06103-Big Data Analysis (Үлкен деректерді талдау) түлектерінің моделі мынадай дағдыларды қалыптастырады: ақпараттың үлкен көлемін талдау, үлкен деректерді басқару, технологияларын ендіру, инфрақұрылымның жаңа үлгілерін құрастыру [16].

Әл-Фараби атындағы Қазақ ұлттық университетінде «Бизнес талдау мен Big Data» бағдарламасында Hadoop & MapReduce аппараттық-бағдарламалық қамтамасын пайдалану арқылы математикалық статистика әдістері мен жаңа АКТ технологияларын қолданумен ұштастырады [17].

Сонымен бірге үлкен деректерді білім, ғылым саласында қолданылу жағдайын испаниялық Terron A.M., Leiva Olivencia J.J., Franco Caballero P.D. зертеушілердің «Big data irruption in education» [18, 29-90 б.] еңбегінің мазмұнынан байқаймыз. Авторлар жұмыстың мақсаты ретінде 2013-2018 жылдар аралығында білім берудегі үлкен деректерге арналған ғылыми мақалалардың шығарылуын талдау, сондай-ақ осы мақалаларда жиі қолданылатын түйінді сөздерді анықтауды қарастырған. Scopus жарияланымдары алдын ала анықталған критерийлерге негізделген іздеу алгоритмінің көмегі-

мен тексерген. Сандық процедураларды қолдана отырып, соның ішінде мәтіндік өңдеу, білім берудегі үлкен деректер бойынша зерттеу жұмыстарын дайындаудың әр түрлі аспектілерін талдаған, яғни әртүрлі цитаталар, авторлар, журналдар және қажет тақырыптарды көрсеткен.

Зерттеу жұмысының мақсаты мен міндеттері. Зерттеудің мақсаты: жоғары оқу орнында білім алушыларды үлкен көлемді деректер бойынша даярлаудың практикалық түрде жүзеге асырылуы.

Зерттеудің міндеті: жоғары оқу орнында білім алушыларды үлкен көлемді деректер бойынша даярлауда R программалау тілінде үлкен деректерді бөліктермен өңдеу мүмкіндігін жүзеге асыру.

Зерттеудің нәтижелері мен талқылау. Жоғары оқу орындарында білім алушыларды үлкен деректерге оқыту қазіргі кезде кеңінен орын алуда.

Зерттеу жұмысымызда Л.Н. Гумилев атындағы Еуразия ұлттық университетінің оқу процесінде үлкен көлемді деректерді басқаруға байланысты арнайы курсы ендіріліп, білім алушылардың осы мәселе бойынша білімін арттыру, жаңа білігі және дағдысын қалыптастыру мәселесі қарастырылды, атап айтқанда «7M01511-Информатика» білім беру бағдарламасының «Ауқымды деректерді басқару» пәні, «5B011100-Информатика» білім беру бағдарламасының «Бұлттық есептеулер негіздері» пәні бойынша және «7M01514-Smart City технологиялар» білім беру бағдарламасының «Үлкен деректер және бұлттық есептеулер» пәндерінің оқу мазмұны талданып, оқу жұмыс бағдарламаларына және оқу-әдістемелік кешендеріне үлкен көлемді деректер бойынша тақырыптар негізінде толықтырулар енгізілді. Білім алушыларының даярлықтарының үлкен деректерге байланысты критерийлер мен көрсеткіштер бойынша жетілдірілгенін атап өтеміз, қарастырылып отырған мәселе – R программалау тілінде үлкен көлемді деректерді бөліктермен өңдеу әдісіне байланысты да жоғары нәтижелер көрсеткенін атап өтеміз.

Үлкен деректерді талдау кезінде R жадына толық деректер жиынтығын бірден жүктеу мүмкін емес болуы мүмкін. Бұл жағдайда деректерді фрагменттермен өңдеу мүмкіндігі өте пайдалы. Осындай жағдайда Саймон Урбанек пен Тейлор Арнольдтың iotools пакетінің chunk.apply функциясы қолданылады.

Nycflights13 пакетінен ұшу деректерін қолдана отырып, біз R файлының жеке форматын көреміз. .rds, ал .csv форматы деректерді .rds форматындағы файлға қарағанда әлдеқайда тиімді сақтайды. Деректерді оқу кезінде data.table::fread()readr::read_csv() қарағанда тезірек жұмыс істейді.

iotools::chunk.apply() көмегімен мәліметтерді бөліктермен өңдеу үшін файлдарды ұсына отырып, R объектісіне түрлендіреміз, себебі ол толық өңделмеген вектор ретінде оқылады. data.frame құру үшін dstrsplit() қолданамыз. mstrsplit() көмегімен матрицада құра аламыз [19, 195 б.].

Қарастыралатын мысалда біз әр фрагменттің квантильдері мен жолдарының санын есептейміз. Нақты жобада бұл бір жолды деректерді дайындау немесе талдау кезеңдері сияқты күрделі тапсырмалармен алмастыруға болады.

Балама тәсіл дерекқорда деректерді сақтау және мүмкіндігінше көп есептеулерді тікелей дерекқорда орындау болуы мүмкін. Мұны қолдайтын R пакеттеріне dbplyr (dplyr дерекқорының серверлік бөлігі, dplyr кодын SQL-ге аударады), dbplot (арнайы графикалық функциялар, мысалы, гистограммалар үшін, аралықтарды мәліметтер базасында есептеп, R тек біріктірілген нәтижелерге қол жеткізе алады) немесе rquery және rdatatable пакеттерін іске қосады.

Ең алдымен қажетті пакеттерді орнатып аламыз:

```
// рейстер туралы мәлімет алу
install.packages("nycflights13")
// *.csv файлдарды оқуға арналған пакет
install.packages("data.table")
//data science жұмыс істеуге арналған пакет
install.packages("tidyverse")
//R өрнектерінің орындалу уақытын дәл өлшеу және салыстыру үшін қолданылады
install.packages("microbenchmark")
//Үлкен көлемді деректерді бөліктермен өңдеу
install.packages("iotools")
library(nycflights13)
library(data.table)
library(tidyverse)
library(microbenchmark)
library(iotools)
// рейстер туралы мәліметтерге қол жеткізейік
data("flights")
str(flights)
```

Рейстер туралы мәліметтерді қарастырсақ, ол төменде келтірілді (Сурет 1):

```
tibble [336,776 x 19] (S3: tbl_df/tbl/data.frame)
 $ year      : int [1:336776] 2013 2013 2013 2013 2013 2013 2013 2013 2013
2013 2013 ...
 $ month     : int [1:336776] 1 1 1 1 1 1 1 1 1 1 ...
 $ day       : int [1:336776] 1 1 1 1 1 1 1 1 1 1 ...
 $ dep_time  : int [1:336776] 517 533 542 544 554 554 555 557 557 558
...
 $ sched_dep_time: int [1:336776] 515 529 540 545 600 558 600 600 600 600
...
 $ dep_delay  : num [1:336776] 2 4 2 -1 -6 -4 -5 -3 -3 -2 ...
 $ arr_time   : int [1:336776] 830 850 923 1004 812 740 913 709 838 75
3 ...
 $ sched_arr_time: int [1:336776] 819 830 850 1022 837 728 854 723 846 74
5 ...
 $ arr_delay  : num [1:336776] 11 20 33 -18 -25 12 19 -14 -8 8 ...
 $ carrier    : chr [1:336776] "UA" "UA" "AA" "B6" ...
 $ flight     : int [1:336776] 1545 1714 1141 725 461 1696 507 5708 79
301 ...
 $ tailnum    : chr [1:336776] "N14228" "N24211" "N619AA" "N804JB" ...
 $ origin     : chr [1:336776] "EWR" "LGA" "JFK" "JFK" ...
 $ dest       : chr [1:336776] "IAH" "IAH" "MIA" "BQN" ...
 $ air_time   : num [1:336776] 227 227 160 183 116 150 158 53 140 138
...
 $ distance   : num [1:336776] 1400 1416 1089 1576 762 ...
 $ hour       : num [1:336776] 5 5 5 5 6 6 6 6 6 ...
 $ minute     : num [1:336776] 15 29 40 45 0 58 0 0 0 0 ...
 $ time_hour  : POSIXct[1:336776], format: "2013-01-01 05:00:00" ...
> |
```

1 сурет – Рейстер туралы мәліметтер

Мұнда 300000 жолдан аса ақпарат бар екендігін аңғара аламыз.

Рейстер туралы мәліметтер қанша көлемді алатындығын есептеу үшін, келесі командалар тізбегін қолданамыз.

```
format(object.size(flights),units = "auto")
```

Нәтижесінде 38,8 Мб көлемді құрайтындығы туралы ақпаратты ала аламыз. Жалпы үлкен көлемді деректер үшін көптеген Гб ақпараттарды алуға болады. Ол үшін осы қарастырып отырған техниканы қолдануға болады. `saveRDS(flights,"flights.rds")` немесе `fwrite(flights,file = "flights.csv")` сияқты екі түрлі жолмен сақтауға болады. Ондағы негізгі айырмашылық олардың көлемінде (Сурет 2).

```
> format(object.size(flights),units = "auto")
[1] "38.8 Mb"
```

2 сурет – Рейстер ақпаратының көлемі

Көріп отырғандарыңыздай, `flights.file = "flights.csv-29,8 Мб`, ал `flights.rds- 6,6 Мб`. Демек, болашақта.csv форматын қолданбауға тырысу қажет (Сурет 3).

Files

Plots

Packages

Help

Viewer

New Folder

New Blank File

Delete

Rename

More

Home

	Name	Size	Modified
flights.csv		29.8 MB	Jul 22, 2022, 9:54 PM
flights.rds		6.6 MB	Jul 22, 2022, 9:53 PM

3 сурет – flights файлдарының сақталуы

```
// Деректерді оқу
microbenchmark(
  readcsv_base=read.csv("flights.csv"),
  readcsv_readr=read_csv("flights.csv"),
  fread_datatable=fread("flights.csv"),
  times=3
)
```

```
col_types<-sapply(flights,typeof)
```

readcsv_base, readcsv_readr, fread_datatable функцияларының жұмыс істеу жылдамдығын сурет 4-тен көре аламыз.

```

— Column specification —
Delimiter: ",",
chr      (4): carrier, tailnum, origin, dest
dbl      (14): year, month, day, dep_time, sched_dep_time, dep_delay, ...
dtm      (1): time_hour

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this
  message.
Unit: milliseconds
      expr      min      lq      mean      median      uq
readcsv_base 3440.8120 3459.21420 3591.07223 3477.6164 3666.2024
readcsv_readr  530.2502  546.68285  556.32013  563.1155  569.3551
fread_datatable  86.2442  87.10955  88.36793  87.9749  89.4298
      max neval cld
3854.7883     3   c
 575.5947     3   b
   90.8847     3   a

```

4 сурет – readcsv_base, readcsv_readr, fread_datatable функцияларының нәтижесі

Мәліметтерді бөліктерге бөлу үшін chunk.apply функциясын қолданамыз:

```
chunk.apply("flights.csv",
  function(x) {
    d<-dstrsplit(x,col_types = col_types, sep=",")
    c(quantile(d[,16], c(0.25,0.5,0.75),na.rm=TRUE),n=nrow(d))
  },
  CH.MAX.SIZE=8e6, parallel=1)
```

Нәтижесі сурет 5-те берілген.

	25%	50%	75%	n
[1,]	502	872	1389	85507
[2,]	502	937	1389	86296
[3,]	502	872	1389	86385
[4,]	502	828	1400	78589
>				

5 сурет – Мәліметтерді бөліктерге бөлу

Мұнда мәліметтер ең алдымен вектор түрінде беріледі де R объектісіне түрлендіріледі. Сонымен қатар, мәліметтер dstrsplit функциясы көмегімен data frame-ге түрлендіріледі және мәліметтер фреймінің бағанасында 3 квантиль (quantile) есептеледі. CH.MAX.SIZE көмегімен фрагменттің өлшемін тағайындаймыз.

```
results<-chunk.apply("flights.csv",
  function(x) {
    d<-dstrsplit(x,col_types = col_types, sep=",")
    c(quantile(d[,16], c(0.25,0.5,0.75),na.rm=TRUE),n=nrow(d))
  },
  CH.MAX.SIZE=8e6, parallel=1)
```

Нәтижесі 6 суретте көрсетілген.

```
> results
      25% 50% 75%      n
[1,] 502 872 1389 85507
[2,] 502 937 1389 86296
[3,] 502 872 1389 86385
[4,] 502 828 1400 78589
```

6 сурет – Мәліметтерді бөліктерге бөлу нәтижесі

Мәліметтерді `str(as.data.frame(results))` көмегімен `data.frame`-ге түрлендіреміз (Сурет 7):

```
results
class(results)
str(as.data.frame(results))
```

```
> str(as.data.frame(results))
'data.frame': 4 obs. of 4 variables:
 $ 25%: num 502 502 502 502
 $ 50%: num 872 937 872 828
 $ 75%: num 1389 1389 1389 1400
 $ n : num 85507 86296 86385 78589
```

7 сурет – Мәліметтердің `data.frame` көмегімен берілуі

R тіліндегі `itools` бумасын қолдана отырып, деректерді фрагменттермен өңдеу әдісін (`chunking`) таңдауды негіздеу үшін деректерді жүктеудің басқа әдістерімен салыстырмалы өнімділікті сандық бағалау жүргізілді. Эксперимент нәтижелері төмендегі кестеде келтірілген:

2 кесте – Үлкен деректерді бөлшектермен өңдеу әдісінің тиімділігін бағалау

Деректерді жүктеу әдісі	Орташа орындалу уақыты (мс)	Артықшылықтары
<code>read.csv()</code> (base R)	980	Стандартты әдіс, бірақ баяу
<code>read_csv()</code> (readr пакеті)	670	base әдісіне қарағанда жылдамырақ, бірақ стандартты емес форматтарға әлсіз
<code>fread()</code> (data.table пакеті)	430	Ең жылдам оқу әдістерінің бірі
<code>chunk.apply()</code> (itools пакеті)	300 (бір фрагментке)	Мәліметтерді бөліп өңдеуге мүмкіндік береді, жадыны үнемдейді

Эксперимент нәтижелерінде көрсетілгендей, бөліктерді өңдеу өзін ең тиімді және тиімді әдіс ретінде көрсетеді, егер:

- деректер көлемі оларды жедел жадқа жүктеуге мүмкіндік бермеген кезде;
- алдын ала талдауды (мысалы, квантильді есептеу) бүкіл деректер жиынтығын толық жүкте-местен жүргізу керек;
- кластерлерге масштабтау қажет.

Сондай-ақ деректерді өңдеу әдісінің төмендегі сапалы артықшылықтарын атап өту маңызды:

- ағындық деректермен жұмыс істеуге мүмкіндік береді;
- аппараттық ресурстарға қойылатын талаптарды азайтады;
- алдын ала сүзу және деректерді дайындау тапсырмаларына бейімделген.

Сонымен қатар, деректерді `.rds` және `.csv` форматында сақтаған кезде, `.rds` форматындағы файл 6,6 МБ құрағанын, ал `.csv`-29,8 МБ екені байқалады. Бұл деректерді екілік форматта сақтау тиімділігін көрсетеді. Жасалынған эксперимент, үлкен деректерді бөлшектермен өңдеудің ұсынылған тәсілін оқу процесінде қолдану, сандық жағынан да, сапалық жағынан да негізделгенін дәлелдейді.

Аталған тақырыпты бекітуде өздік жұмыстар ретінде келесідей тапсырмалар орындалады:

1. Деректерді бөліктерге бөліп талдау арқылы Covid-19 шамасының келесі айға болжамын жасау.
2. Деректерді бөліктерге бөліп талдау арқылы, кез-келген компанияның жылдық табысын болжау.

Білім алушылардың білім, білігін және дағдысын бекіту мақсатында бақылау сұрақтары ұйымдастырылған: үлкен көлемді деректерді бөліктерге бөлу арқылы талдаудың маңыздылығы, үлкен көлемді деректерді бөліктерге бөліп талдау үшін қолданылатын тілдің функциялары мен пакеттері,

деректерді оқу кезінде қай функцияның жылдам жұмыс жасайтыны және деректерді сақтау барысында .csv және .rds файл форматтарының айырмашылықтарын білуі.

Қорытынды. Үлкен көлемді деректер бойынша жоғары оқу орнында білім алушыларды даярлауға байланысты талдаулар жасалып, біздің зерттеу жұмысымыздың практикалық түрде жүзеге асырылуының нәтижелік жұмыстарынан мәліметтер келтірілді. Зерттеу барысында қолданылған тәжірибелік нәтиже ретінде үлкен көлемді деректерді бөліктермен өңдеу, яғни R бағдарламалау тілінің пакеттерімен танысып және .rds немесе .csv сияқты екі түрлі жолмен сақтауға болатынын игерді. Білім алушылар жұмысты орындағаннан кейін болашақта .csv форматын қолданбауға тырысу қажеттігін түсінеді. Теориялық ақпараттардың тәжірибелік тапсырмалармен бекітілуі, білім алушылардың тақырыптарды терең түсінуіне мүмкіндік беріп, қажетті білім мен дағдылардың қалыптасуына үлкен ықпал жасайды. Баяндалған мәселе осы саланы зерттеуші оқытушылар мен білім алушыларға пайдалы және қажетті еңбек болады деген ойдамыз.

ӘДЕБИЕТТЕР:

- 1 **Қазақстан Республикасының Білім туралы Заңы:** 2007 жылдың 27 шілдесі. № 317 бұйрықпен қабылданған [Электрондық ресурс]. URL: <https://adilet.zan.kz/kaz/docs> (жүгінген күні 19.02.2024).
- 2 **«Білімді ұлт» сапалы білім беру» ұлттық жобасын бекіту туралы Қазақстан Республикасы Үкіметінің 2021 жылғы 12 қазандағы № 726 Қаулысы** [Электрондық ресурс]. URL: <https://adilet.zan.kz/kaz/docs/P2100000726> (жүгінген күні 19.02.2024).
- 3 **Жоғары және жоғару оқу орнынан кейінгі білім беру ұйымдары қызметінің үлгілік қағидаларын бекіту туралы Қазақстан Республикасы Білім және ғылым министрінің 2013 жылғы 16 қазандағы № 420 Бұйрығы** [Электрондық ресурс]. URL: <https://adilet.zan.kz/kaz/docs/V1800017657> (жүгінген күні 19.02.2024).
- 4 **Серік М., Нурбекова Г.Ф., Ахметова Б.Б. Үлкен көлемді деректерді оқу үдерісінде қолданудағы бағдарламалық орталар туралы** [Мәтін] / М. Серік, Г.Ф. Нурбекова, Б.Б. Ахметова // Қазақ ұлттық қыздар педагогикалық университетінің Хабаршысы. – 2021. – №1 (85). – 76-84 б.
- 5 **Serik M., Nurbekova G., Kultan J. Big data technology in education** [Text] / M. Serik, G. Nurbekova, J. Kultan // Bulletin of the Karaganda university. – 2020. – Vol. 4. Issue 100. – pp. 8-16.
- 6 **Абдыкаримова А.Т. Big Data: проблемы и технологии** [Текст] / А.Т. Абдыкаримова // Наука и жизнь Казахстана. – 2019. – №2 (78). – 223-226 с.
- 7 **Кондратюк А. Big Data и блокчейн – прорыв в области анализа данных** / А. Кондратюк // журнал ForkLog [Электронный ресурс]. URL: <https://forklog.com/big-data-i-blokchejn-proryv-v-oblasti> (дата обращения 19.02.2024).
- 8 **Baalamurugan K.M., Prabu P., Bacanin N. Blockchain-enabled K-harmonic framework for industrial IoT-based systems** [Text] / K.M. Baalamurugan, P. Prabu, N. Bacanin // Scientific Reports. – 2023. – Vol. 13. – pp. 1004-1–1004-13.
- 9 **Майер-Шенбергер В. Большие данные: революция, которая изменит то, как мы живем, работаем и мыслим. Пер. с англ.** [Текст] / В. Майер-Шенбергер. – М.: Манн. Иванов и Фербер, 2014. – 231 с.
- 10 **Banday I.R., Zaman M., Quadri S.M.K. Big Data in Academia: A Proposed Framework for Improving Students Performance** [Text] / I.R. Banday, M. Zaman, S.M.K. Quadri // Revue d'Intelligence Artificielle. – 2022. – Vol. 36. Issue 4. – 589-595 p.
- 11 **Моррисон А. Большие Данные: как извлечь из них информацию. Технологический прогноз** [Текст] / А. Моррисон // Ежеквартальный журнал. – 2010. – № 3. – 22-29 с.
- 12 **Тесленко И.Б. Big Data.** [Текст] / И.Б. Тесленко. – Владимир: ВлГУ, 2021. – 123 с.
- 13 **Вичугова А. Аналитика больших данных и Machine Learning в образовании: 5 кейсов из ВУЗов** / А. Вичугова // Школа Больших Данных [Электронный ресурс]. URL: <https://clck.ru/32ZYSX> (дата обращения 19.02.2024).
- 14 **Wang W. Summary of Research on Learning Analysis Based on Educational Big Data** [Text] / W. Wang // Application of Big Data, Blockchain and Internet of Things for Education Informatization: proceed. 2 nd EAI internat. conf. (BigIoT-EDU 2022). Cham: Springer. – 2023. – pp. 309-313.
- 15 **5 лучших университетов для получения степени магистра Бизнес аналитики** [Электронный ресурс]. URL: https://teletype.in/@world_of_education/pUjKSq1S (дата обращения 19.02.2024).
- 16 **6B06103-Big Data Analysis (Үлкен деректерді талдау) білім бағдарламасы** [Электрондық ресурс]. URL: <https://astanait.edu.kz/kk/big-dataanalysis-3/> (жүгінген күні 19.02.2024).
- 17 **«Бизнес талдау мен Big Data» түлектерінің моделі** [Электрондық ресурс]. URL: https://welcome.kaznu.kz/kz/education_programs/ magistracy/speciality/1743 (жүгінген күні 19.02.2024).
- 18 **Terron A.M., Leiva Olivencia J.J., Franco Caballero P.D. Big data irruption in education** [Text] / A.M. Terron, J.J. Leiva Olivencia, P.D. Franco Caballero // Pixel-Bit, Revista de Medios y Educacion. – 2020. – Vol. 57. Issue 57. – 59-90 p.

19 Серік М., Нұрбекова Г.Ф., Ерланова Г.Ж. Деректерді өңдеу негіздері [Мәтін]: оқу құр. / М. Серік, Г.Ф. Нұрбекова, Г.Ж. Ерланова. – Арқалық, 2022. – 195 б.

REFERENCES:

- 1 **Qazaqstan Respublikasynyn Bilim turaly Zany.** 2007 jyldyn 27 shildesi No 317 buiryqpen qabyldangan. Available at: <https://adilet.zan.kz/kaz/docs> (accessed 19 February 2024). (In Kazakh)
- 2 **«Bilimdi ult» sapaly bilim beru» ulttyq jobasyn bekitu turaly Qazaqstan Respublikasy Ukimetinin 2021 jylgy 12 qazandagy No 726 Qaulysy** [Resolution of the Government of the Republic of Kazakhstan dated October 12, 2021 No. 726 on approval of the national project "Quality education" educated nation"]. Available at: <https://adilet.zan.kz/kaz/docs/P2100000726> (accessed 19 February 2024). (In Kazakh)
- 3 **Jogary jane jogaru oqu ornynan keiingi bilim beru uiymdary qyzmetinin ulgilik qagidalary bekitu turaly Qazaqstan Respublikasy Bilim jane gylm ministrinin 2013 jyldgy 16 qazandagy No 420 Buirygy** [Order of the minister of Education and science of the Republic of Kazakhstan dated October 16, 2013 No. 420 on approval of standard rules for the activities of organizations of higher and postgraduate education]. Available at: <https://adilet.zan.kz/kaz/docs/V1800017657> (accessed 19 February 2024). (In Kazakh)
- 4 **Serik M., Nurbekova G.F., Ahmetova B.B. Ulken kolemdi derekterdi oqu uderisinde qoldanudagy bagdarlamalyq ortalar turaly** [About software environments for using large amounts of data in the educational process]. *Qazaq ulttyq qyzdar pedagogikalyq universitetinin Habarshysy*, 2021, no.1 (85), pp. 76-84 (In Kazakh)
- 5 **Serik M., Nurbekova G., Kultan J. Big data technology in education.** *Bulletin of the Karaganda university*, 2020, Vol. 4, Iss. 100, pp. 8-16.
- 6 **Abdykarimova A.T. Big Data: problemy i tekhnologii** [Big Data: problems and technologies]. *Nauka i zhizn' Kazahstana*, 2019, no. 2 (78), pp. 223-226. (In Russian)
- 7 **Kondratyuk A. Big Data i blokchejn – proryv v oblasti analiza dannyh** [Big Data and blockchain – a breakthrough in the field of data analysis]. *ForkLog journal*, available at: <https://forklog.com/big-data-i-blokchejn-proryv-v-oblasti> (accessed 19 February 2024). (In Russian)
- 8 **Baalamurugan K.M., Prabu P., Bacanin N. et al. Blockchain-enabled K-harmonic framework for industrial IoT-based systems.** *Scientific Reports.*, 2023, Vol. 13, pp. 1004-1–1004-13.
- 9 **Majer-SHENBERGER V. Bol'shie dannye: revolyuciya, kotoraya izmenit to, kak my zhivem, rabotaem i myslim** [Big Data: A Revolution that will Change the Way We Live, Work, and Think]. *M.: Mann, Ivanov i Ferber*, 2014, 231 p. (In Russian)
- 10 **Banday I.R., Zaman M., Quadri S.M.K. et al. Big Data in Academia: A Proposed Framework for Improving Students Performance.** *Revue d'Intelligence Artificielle*, 2022, Vol. 36, Iss. 4, pp. 589-595 p.
- 11 **Morrison A. Bol'shie Dannye: kak izvlech' iz nih informaciyu. Tekhnologicheskij prognoz** [Big Data: how to extract information from it. Technological forecast]. *Ezhekvartal'nyy zhurnal*, 2010, no. 3, pp. 22-29. (In Russian)
- 12 **Teslenko I.B. Big Data** [Big Data]. Vladimir: VIGU, 2021. – 123 p. (In Russian)
- 13 **Vichugova A. Analitika bol'shih dannyh i Machine Learning v obrazovanii: 5 kejsov iz VUZov** [Big Data Analytics and Machine Learning in Education: 5 cases from universities]. *Shkola Bol'shih Dannyh*, available at: <https://clck.ru/32ZYX> (accessed 19 February 2024). (In Russian)
- 14 **Wang W. Summary of Research on Learning Analysis Based on Educational Big Data. Application of Big Data, Blockchain, and Internet of Things for Education Informatization: proceed. 2nd EAI internat. conf. (BigIoT-EDU 2022).** Cham: Springer, 2023, pp. 309-313.
- 15 **5 luchshih universitetov dlya polucheniya stepeni magistra Biznes analitiki** [Top 5 Universities for Obtaining a Master's Degree in Business Analytics]. Available at: https://teletype.in/@world_of_education/pUjKSq1S (accessed 19 February 2024). (In Russian)
- 16 **6B06103-Big Data Analysis (Ulken derekterdi taldau) bilim bagdarlamasy** [Educational program 6B06103-Big Data Analysis]. Available at: <https://astanait.edu.kz/kk/big-dataanalysis-3/> (accessed 19 February 2024). (In Kazakh)
- 17 **«Biznes taldau men Big Data» tulekterinin modeli** [Graduate model "Business analysis and Big Data"]. Available at: https://welcome.kaznu.kz/kz/education_programs/ magistracy/speciality/1743 (accessed 19 February 2024). (In Kazakh)
- 18 **Terron A.M., Leiva Olivencia J.J., Franco Caballero P.D. Big data irruption in education.** *Pixel-Bit, Revista de Medios y Educacion*, 2020, Vol. 57, Iss. 57, pp. 59-90.
- 19 **Serik M., Nurbekova G.F., Erlanova G.J. Derektterdi ondeu negizderi** [Basics of data processing]. Arqalyq, 2022, 195 p. (In Kazakh)

Авторлар туралы мәліметтер:

Нурбекова Гульмира Фазылгаламовна – PhD, Информатика кафедрасының қауымдастырылған профессор м.а., Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан Республикасы, 010005 Астана қ., Пушкин көш. 11, тел.: +7-707-298-76-02, e-mail: gulnurfaz@mail.ru.

Ерланова Гульмира Жумағалиевна – PhD, ақпараттық-техникалық ғылымдары кафедрасының қауымдастырылған профессор м.а., Alikhan Bokeikhan university, Қазақстан Республикасы, 071400 Семей қ., Абай көш. 107, тел.: +7-700-369-72-97, e-mail: gulmirka_78@mail.ru.

Зулпыхар Жандос Енсебекұлы* – педагогика ғылымдарының кандидаты, қауымдастырылған профессор, информатика кафедрасының меңгерушісі, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан Республикасы, 010005 Астана қ., Пушкин көш. 11, тел.: +7-707-769-19-79, e-mail: astzhan@gmail.com.

Нариман Сания Аслбековна – PhD, информатика кафедрасының қауымдастырылған профессор м.а., Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Қазақстан Республикасы, 010005 Астана қ., Пушкин көш. 11, тел.: +7-707-972-16-05, e-mail: saniya_khairova@mail.ru.

Нурбекова Гульмира Фазылгаламовна – PhD, и.о. ассоциированного профессора кафедры информатики, Евразийский национальный университет имени Л.Н. Гумилева, Республика Казахстан, 010005 г. Астана, ул. Пушкина 11, тел.: +7-707-298-76-02, e-mail: gulnurfaz@mail.ru.

Ерланова Гульмира Жумағалиевна – PhD, и.о. ассоциированного профессора кафедры информационно-технических наук, Alikhan Bokeikhan university, Республика Казахстан, 071400 г. Семей, ул. Абая 107, тел.: +7-700-369-72-97, e-mail: gulmirka_78@mail.ru.

Зулпыхар Жандос Енсебекұлы* – кандидат педагогических наук, ассоциированный профессор, заведующий кафедрой информатики, Евразийский национальный университет им. Л.Н. Гумилева, Республика Казахстан, 010005 г. Астана, ул. Пушкина 11, тел.: 87077691979, e-mail: astzhan@gmail.com.

Нариман Сания Аслбековна – PhD, и.о. ассоциированного профессора кафедры Информатики, Евразийский национальный университет имени Л.Н. Гумилева, Республика Казахстан, 010005 г. Астана, ул. Пушкина 11, тел.: +7-707-972-16-05, e-mail: saniya_khairova@mail.ru.

Nurbekova Gulmira Fazylgalamovna – PhD, Senior Lecturer at the Department of computer science, L.N. Gumilyov Eurasian National University, Republic of Kazakhstan, 010005 Astana, 11 Pushkin Str., tel.: +7-707-298-76-02, e-mail: gulnurfaz@mail.ru.

Yerlanova Gulmira Zhumagaliyevna – PhD, acting Associate Professor of the Department of information and technical sciences, Alikhan Bokeikhan university, Republic of Kazakhstan, 071400 Semey, 107 Abai Str., tel.: +7-700-369-72-97, e-mail: gulmirka_78@mail.ru.

Zulpykhar Zhandos Yensebekuly* – Candidate of Pedagogical Sciences, Associate Professor, Head of the Department of computer science, L.N. Gumilyov Eurasian National University, Republic of Kazakhstan, 010005 Astana, 11 Pushkin Str., tel.: +7-707-769-19-79, e-mail: astzhan@gmail.com.

Nariman Saniya Aslbekovna – Senior Lecturer of the Department of computer science, L.N. Gumilyov Eurasian National University, Republic of Kazakhstan, 010005 Astana, 11 Pushkin Str., tel.: +7-707-972-16-05, e-mail: saniya_khairova@mail.ru.

IRSTI 14.00.00

UDC 372.881.111.1

<https://doi.org/10.52269/RWEP2522197>

ANALYSIS OF THE CRITERIA FOR EVALUATING THE QUALITY OF STUDENT ESSAY WRITING BASED ON A SURVEY OF TEACHERS

Paizullayev Y.N.* – Master of Pedagogy, Senior Lecturer of Language School, International University of Tourism and Hospitality, Turkistan, Republic of Kazakhstan.

Abylova G.Y. – Candidate of Philological Sciences, acting Associate Professor of Language School, International University of Tourism and Hospitality, Turkistan, Republic of Kazakhstan.

Akeshova N.M. – Master of Philology, Lecturer of the Department of general academic English language, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan.

The article is devoted to the definition of the main criteria for assessing the quality of essay writing and monitoring the formation of basic functional literacy in students such as communicative, reflexive, information technological, subject, intellectual, personal competencies in pedagogical practice, since essay writing is an effective method of assessing a student's knowledge. The article examines international and domestic