С.А. „Д. А. Ценов" - Свищов
Вх.№ 93.00-90/03.11.2017 г.

Р Е Ц Е Н З И Я

на дисертационен труд за придобиване на образователната и научна степен „доктор"

Рецензент: **проф. д-р Васил Мишов Цанов**, в област "Социални науки, стопански и правни науки", научна специалност „Статистика и демография" и "Народно стопанство (вкл. регионална икономика и стопанска история)", заемащ длъжността "професор" в секция "Макроикономика" в Института за икономически изследвания на БАН

Автор на дисертационния труд: **Галя Живкова Статева**, докторант в задочна докторантура по научната специалност „Статистика и демография" към катедра „Математика и статистика" при СА „Д. А. Ценов" – Свищов

Тема на дисертационния труд: **"Проблеми при статистическия анализ на „Големи данни" (BIG DATA)"**

Основание за представяне на рецензията: участие в състава на научното жури по защита на дисертационния труд съгласно заповед № 950/05.10.2017 г. на Ректора на СА „Д. А. Ценов" – Свищов и решение на научното жури на своето първо заседание от 12.10. 2017 г.

I. Общо представяне на дисертационния труд

Обект, предмет, цели и задачи на дисертационния труд. В дисертационния труд предметът и обектът на изследването са ясно очертани. Обект на изслеването са Big Data, а предметът е фокусиран върху възможностите за тяхното използване за обогатяване (допълване) на съществуващата статистическа информация. Целта на дирсертационния труд е ясно и точно формулирана, а именно да се изследват теоретико-методологичните и практически аспекти на Big Data с оглед разработването на цялостна концепция за тяхното приложение в българската статистическа практика. За постигането на тази цел са дефинирани 7 основни задачи, които напълно съответстват на поставената цел. Изследователската теза е изчерпателно представена. Тя съответства на съдържанието на дисертационния труд и в хода на изследването е доказана.

Обем и структура на дисертационния труд. Дисертационният труд е в обем от 208 стандартни страници, от които 178 страници основен текст, 15 страници библиографска справка и 15 страници приложения. Основният текст включва 8 таблици

и 16 фигури. В структурно отношение се състои от увод, две глави, заключение, използвана литература и приложения.

Първа глава има подчертан теоретико-методологически характер. Посветена е на проблемите, свързани с дефинирането на Big Data и тяхното използване в официалната статистическа практика. В нея изчерпателно и задълбочено се изследват няколко ключови проблема, отнасящи се до: а) същността и източниците на Big Data; б) необходимостта от усъвършенстване на концепцията на класическото статистическо изследване с оглед използването на Big Data; в) оценка на качеството на големите данни; г) технологичните и методически проблеми при обработката на Big Data.

Втора глава съдържа емпирично изследване на приложението на големите данни в статистическата практика. В съдържателно отношение изследването представалява извличане на информация от интернет за характеристиките на предприятията в страната. Изчерпателно са представени концептуалната рамка на емпиричното изследване, технологията за извличане на информация и е направен сравнителен анализ с резултатите от регулярното статистическо изследване на НСИ „Използване на ИКТ (информационни и комуникационни технологии) в предприятията". Направени са важни изводи на възможността за приложените на изполдвания инструментариум за набиране, структуриране, обработка, съхранение и анализ на получената информация.

Използвана литература. Използваната литература е в обем от 148 заглавия, от които 39 заглавия на български и 109 заглавия на латиница. Тя обхваща съвременните достижения и емпирични изследвания в дадената област. Това свиделства за добрата информираност на дисертанта по изследваната проблематика.

Процедурните изисквания за допускане до защита са спазени. Дисертационният труд е обсъден и насочен за защита от катедра "Математика и статистика" към факултет „Стопанска отчетност" на СА „Д. А. Ценов" - Свищов.

II. Преценка на формата и съдържанието на дисертационния труд

Актуалност на научната проблематика. Дисертационният труд е посветен на съвременна тема, която е обект на значими теоретико-методологически и приложни изследвания в целия свят. Изследваната проблематика е изключително актуална за осъвременяването на статистическите изследвания. Тя е особено актуална и за България не само с оглед хармонизирането на статистическата информация с международните и европейски стандарти, но и с решаването на редица практически

въпроси, отнасящи се до използването на огромните масиви от данни в интернет пространството. Поради тази причина тематиката на дисертационния труд е от съществено значение за развитието на методологията на статистическите изследвания в НСИ.

Оценка на разработеността на изследваната проблематика от други учени и изразяване на собствено мнение. Темата за големите числа и тяхното използване в статистическата практика е сравнително нова в световен мащаб. Тя е в процес на развитие както в теоретико-методологически, така и в практически аспект. В България изследваната проблематика е срабвително слабо разработева. Налице са ограничен брой разработки, които дисертантът коректно цитира.

В дисертационния труд се съдържат множество убедителни доказателства за много добрата осведоменост на дисертанта по изследваната проблематика. Основание за това ми дават не само изразените познания за теорията, методологията и инструментариума на Big Data, но и оценките, които авторът дава на съществуващите концепции за приложение на големите данни в статистическата практика. Освен това дисертантът участва в множество национални и международни проекти по проблемите на големите данни.

В дисертационния труд мнението на докторанта по редица въпроси е ясно и аргументирано представено. По всички дискусионни въпроси дисертантът изразява собствено становище. Общото впечатление е, че по редица спорни въпроси авторът изразява собствено мнение.

Оценка на обема на дисертационния труд. Структурата и обемът на дисертационния труд са добре балансирани и напълно съответстват на изследваната проблематика. Структурата е логически издържана и последователна. Тя следва дефинираните проблеми и е подчинена на реализиране на поставената цел и задачи.

Оценка на качеството и броя на предложения илюстративен материал. Представеният илюстративен материал (8 таблици и 16 графики) е добре разработен и представителен. Представените графики и таблици в текста са достатъчно информативни и разбираеми. Част от тях са взаимствани от чужди публикации, които коректно са цитирани.

Оценка на стила и начина на представяне. Дисертационният труд е написан в научен и стегнат стил, типичен за изследваната проблематика. Стилът е ясен, точен и научно издържан. Използвани са множество специализирани научни термини, които

коректно са преведени на български език. Не съм забелязъл логически повторения и дублиране на текстове.

Авторефератът напълно отразява съдържанието на дисертационния труд. Той представя в синтезиран вид постигнатите от дисертанта резултати.

Оценка на публикациите по дисертацията. По темата на дисертационния труд са направени 7 публикации – една студия, 4 статии и два доклада. Студията е публикувана в Годишен алманах. Научни изследвания на докторанти, през 2016 г. Статиите са публикувани в сп. Статистика и специализирано научно издание на СА. Докладите са самостоятелни и са отпечатани в специализирани научни издания на СА и УНСС. Публикационната дейност на докторанта е напълно достатъчна и отговаря на изискванията за придобиване на ОНС "доктор".

III. Научни и научно-приложни приноси на дисертационния труд

В дисертационния труд са постигнати научни и научно-приложни резултати, които имат характер на доразвитие на съществуващите знания за приложението на Big Data в статистическите изследвания. Те могат да се обобщят в следните.

Първо, представено е авторово виждане за дефинирането и основните характеристики на големите данни с оглед тяхното приложение в статистическите изследвания. Считам, че представената дефиниция е обобщен израз на съществуващите виждания, пречупени през призмата на тяхното използване за обогатяване на съществуващата статистическа информация.

Второ, направена е систематизация и оценка на основните източници на Big Data с оглед тяхното приложение в статистическата практика. На тази основа са оценени съответствията (респ. несъответствията) между характеристиките на големите данни и тези на традиционните източници на данни. В резултат на този сравнителен анализ е направен важния извод, че големите данни са твърде различни от традиционните и трудно могат да се прилагат в съчетание с тях без допълнителна обработка.

Трето, направени са оценки на качеството на Big Data и предложения за неговото повишаване с цел използването им в статистическата практика. Оценките се отнасят предимно до подходите за оценява и в по-малка степен до методите за оценяване. Подходите за повишаване качеството на големите данни са дискутирани в светлината на тяхното приложение, а именно дали да бъдат използвани като самостоятелен източник на данни или в комбинация с традиционните статистически

изследвания. И в двата случая са налице сериозни предизвикателства пред националните статистики.

Четвърто, проведеното емпирично изследване за събиране на интернет информация за характеристиките на предприятията в страната има иновативен характер за националната статистика. Макар, че то е резултат на екип от експерти, ролята на дисертанта е от съществено значение. На базата на получените резултати са формулирани основните предизвикателства пред националната статистика, отнасящи се до използването на техниките за извличане на информация от интернет пространството.

В дисертационния труд се съдържат приноси, които могат да се обобщят в следните три направления: 1) обогатяване на съществуващите знания с нови виждания и предложения за преодоляване на съществуващите информационни проблеми при използването на Big Data; 2) апробация на съществуващия методологически инструментариум за извличане на информация от интернет пространството за нуждите на статистическата практика; 3) обогатяване на съществуващите познания за практическото приложение на големите данни в националната статистика.

Постигнатите в дисертационния труд научни и научно-приложни приноси считам, че отговарят на изискванията на нормативната уредба за получаването на образователната и научна степен "доктор".

IV. Критични бележки, въпроси и препоръки по дисертационния труд

Независимо от постигнатите в дисертационния труд резултати могат да се направят критични бележки и препоръки. Те се изразяват в следното:

1. Анализът на качеството на Big Data се нуждае от по-голяма задълбоченост. Подходите и методите за оценка са твърде общо представени. Това се отнася предимно до методите за оценка. От изложението не става ясно кои конвенционални методи за оценка на качеството могат (или трябва) да бъдат прилагани (или задължителни) за оценка на големите данни с цел тяхното прилагане в статистическата практика.

2. В анализа на палучените резултати от емпиричното изследва липсват оценки за качеството на получените данни. И в четирите изследвани случая са направени констатции за качеството на използвания инструментариум, но по отношение на качеството на данните такива липсват. Например при оценката на случай-2 се твърди, че „използваните ИТ средства и инструменти са достатъчно ефективни (стр. 160)" ,

докато данните от табл. 2.3 не потвърждават тази констатация по отношение на данните.

Въпроси към дисертанта:

1. В каква степен традиционните статистически методи за оценка на качеството на конвенционалните статистически данни могат да се прилагат за оценка на качеството на информацията, получена от Big Data? Не трябва ли те да бъдат едни и същи?

2. Възможно ли е в статистическата практика да се комбинират (допълват) конвенционалните статистически данни с данни от Big Data с непроверено качество? И има ли смисъл от това?

Тези критични бележки би следвало да се възприемат като препоръки за бъдещата научноизследователска дейност в тази област. Те не омаловажават постигнатите научни и научно-приложни резултати в дисертацията.

V. Обобщено заключение и становище

Дисертационният труд представлява едно задълбочено научно изследване, което съдържа необходимите качества за получаването на ОНС „доктор". В него са постигнати научни и научно-приложни резултати, които **напълно съответстват** на изискванията на Закона за развитието на академичния състав в Република България, Правилника за неговото прилагане и Правилника за придобиване на научни степени и заемане на академични длъжности в СА „Д.А.Ценов" - Свищов за получаването на образователната и научна степен "доктор". Изразявам моето **положително становище** и препоръчвам на Научното жури **да присъди** ОНС "доктор" на **Галя Живкова Статева** по научната специалност „Статистика и демография".

08. 11. 2017 г.

Рецензент:
(проф. д-р В. Цанов)