

*Межов Илья Сергеевич, студент 4 курса, направление подготовки «Бизнес-информатика», Финансовый университет при Правительстве Российской Федерации (Липецкий филиал), Липецк, Россия*

## **АЛГОРИТМ ПРОВЕДЕНИЯ СОВМЕСТНОГО АНАЛИЗА НА ПРИМЕРЕ ДЕЯТЕЛЬНОСТИ ПАО «СБЕРБАНК» В КОСТРОМСКОЙ ОБЛАСТИ**

**Аннотация:** В данной статье представлен совместный анализ показателей деятельности ПАО «Сбербанк». Поиск зависимостей между показателями деятельности компании является неотъемлемым инструментом системы бизнес-аналитики, позволяющим выявлять, объяснять и использовать различные закономерности для эффективной деятельности организации. Цель исследования — создание алгоритма проведения совместного анализа информации с помощью доступного программного обеспечения. Актуальность данного исследования вызвана потребностью компаний в обработке бизнес-информации. Объектом исследования стал ПАО «Сбербанк». В ходе работы был проведен совместный анализ с помощью математико-статистических методов, выявлены имеющиеся закономерности, приведена их возможная интерпретация и разработан алгоритм, позволяющий осуществлять совместный анализ с помощью доступного программного обеспечения. Исследование может быть полезным для компаний, организаций и предпринимателей, разрабатывающих собственную систему бизнес-аналитики.

**Ключевые слова:** совместный анализ, бизнес-аналитика, бизнес-информация, Костромская область, ПАО «Сбербанк».

**Annotation:** This article presents a joint analysis of the performance indicators of Sberbank PJSC. The search for dependencies between the performance indicators of a company is an integral tool of the business intelligence system that allows you to

identify, explain and use various patterns for the effective operation of an organization. The purpose of the study is to create an algorithm for joint analysis of information using available software. The relevance of this study is due to the need for companies to process business information. The object of the study was Sberbank PJSC. In the course of the work, a joint analysis was carried out using mathematical and statistical methods, the existing patterns were identified, their possible interpretation was given, and an algorithm was developed that allows for joint analysis using available software. The research can be useful for companies, organizations and entrepreneurs developing their own business intelligence system.

**Key words:** joint analysis, business analytics, business information, Kostroma region, PJSC "Sberbank".

Для проведения совместного анализа было решено осуществить отбор информации по всем имеющимся в открытых данных ПАО «Сбербанк» показателям по Костромской области. У разных показателей отличаются периоды, за которые опубликованы наблюдения, поэтому с целью соблюдения условия однородности данных был выбран период с 15.01.2015 г. по 15.11.2018 г., так как в этом периоде имеются данные по всем показателям [3].

Совместный анализ был осуществлен с помощью программного средства «Microsoft Excel».

Первым этапом работы стал отбор необходимых данных с помощью встроенного инструмента «Фильтр». Данные были собраны в таблицу в удобной для дальнейшей работы форме. Фрагмент полученной таблицы в «Microsoft Excel» приведен на рисунке 1.

		Количество заявок на потребительские кредиты	Средняя сумма заявки на потребительский кредит	Количество заявок на ипотечные кредиты	Средняя сумма заявки на ипотечный кредит	Количество новых депозитов	Средняя сумма нового депозита	Средняя зарплата	Средняя пенсия	В среднем руб. на текущем счете на человека	В среднем депозитов в руб. на человека	Средние расходы по картам	Средний чек в формате Фастфуд	Средний чек в формате Ресторан	Средние траты в ресторане	Средние траты в ресторане
	Дата	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
1	15.01.2015	1200	146338	179	874299	12323	172467	12899	14276	11829	218208	3795	292	775	614	1757
2	15.02.2015	1557	144223	196	1009589	13964	144143	17255	12592	12175	216253	3579	279	721	596	1585
3	15.03.2015	1767	149138	307	1061961	14374	155000	17014	13021	11762	217447	4140	280	667	648	1549
4	15.04.2015	2599	151628	381	1107098	11415	157507	19174	15095	12593	217538	3778	269	633	610	1447
5	15.05.2015	2458	139750	320	995712	7809	157306	17232	11567	12261	221069	4062	276	658	639	1485
6	15.06.2015	4822	152967	350	1101445	8818	172114	19985	12981	12737	226051	4346	283	643	647	1462
7	15.07.2015	4875	151202	416	960094	8831	186242	18888	13820	12827	231269	4578	299	662	659	1535
8	15.08.2015	4415	141656	429	1057841	8766	170275	16808	12992	12309	239114	4699	313	701	663	1663
9	15.09.2015	3774	143724	459	1105133	8291	182044	16404	12942	12061	242496	4423	291	639	665	1542
10	15.10.2015	3794	163535	483	1202742	9656	170102	17192	13459	12249	244246	4776	290	642	682	1587

Рисунок 1. Фрагмент данных для совместного анализа

В соответствии с темой исследования в качестве зависимой переменной  $Y$  был выбран показатель «Количество заявок на потребительские кредиты». Таким образом, четырнадцать оставшихся показателей являются объясняющими переменными  $X_1, X_2, X_3 \dots X_{14}$ .

Вторым этапом работы стало построение таблицы коэффициентов парной корреляции с помощью инструмента «Корреляция» надстройки «Анализ данных» программного продукта «Microsoft Excel» [1]. Полученная таблица коэффициентов парной корреляции представлена на рисунке 2.

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
Y	1														
X1	0,382395	1													
X2	0,637709	0,6135	1												
X3	0,463477	0,836477	0,710232	1											
X4	-0,321	-0,0706	-0,26206	-0,06418	1										
X5	0,307265	0,725675	0,549549	0,680131	0,000349	1									
X6	0,484512	0,271092	0,400467	0,322125	0,035749	0,322774	1								
X7	0,320207	0,201176	0,270136	0,185509	0,26592	0,242107	0,800734	1							
X8	0,571347	0,861308	0,649148	0,86224	-0,07913	0,767676	0,458906	0,320712	1						
X9	0,507809	0,911105	0,69217	0,867898	-0,18629	0,806428	0,348761	0,215514	0,926676	1					
X10	0,549871	0,888063	0,717174	0,900506	-0,11149	0,790099	0,396065	0,281038	0,962148	0,958018	1				
X11	0,06504	-0,24184	-0,28551	-0,33604	-0,33257	-0,29583	-0,18475	-0,16668	-0,31508	-0,17574	-0,26416	1			
X12	-0,51401	-0,87444	-0,74213	-0,86459	0,154438	-0,7897	-0,39409	-0,25142	-0,90529	-0,9527	-0,9232	0,354975	1		
X13	0,566257	0,884244	0,715895	0,897481	-0,1525	0,772911	0,351286	0,221752	0,942345	0,953176	0,984288	-0,20018	-0,91118	1	
X14	-0,43409	-0,77172	-0,6988	-0,7893	0,106928	-0,71744	-0,38381	-0,21149	-0,80133	-0,83372	-0,80287	0,492881	0,947864	-0,78223	1

Рисунок 2. Коэффициенты парной корреляции

На основе полученных значений коэффициентов парной корреляции были установлены расчетные значения  $t$ -статистики Стьюдента с помощью следующей формулы [4]:

$$t_{расч} = \frac{r_{xy} \sqrt{n-2}}{\sqrt{1-r_{xy}^2}},$$

где  $r_{xy}$  — соответствующий коэффициент парной корреляции;

$n$  — количество наблюдений соответствующего показателя.

С помощью функции «СТЮДЕНТ.ОБР.2Х» было получено табличное значение критерия Стьюдента для  $n = 47$ :

$$t_{табл} = 2,014.$$

С помощью функции «ЕСЛИ» было осуществлено сравнению расчетных значений t-критерия с табличным. В случае, если расчетное значение удовлетворяет условию  $t_{расч} > t_{табл}$ , критерий отмечается единицей, в противном случае — нулем. Реализация в программном продукте «Microsoft Excel» представлена на рисунке 12.

Критерий	Значение	Условие
трасч1	2,776	1
трасч2	5,554	1
трасч3	3,509	1
трасч4	-2,274	0
трасч5	2,166	1
трасч6	3,715	1
трасч7	2,267	1
трасч8	4,670	1
трасч9	3,954	1
трасч10	4,416	1
трасч11	0,437	0
трасч12	-4,020	0
трасч13	4,609	1
трасч14	-3,232	0
<b>табл</b>	2,014	

Рисунок 3. Проверка t-критерия

Таким образом, после оценки значимости коэффициентов корреляции с помощью t-статистики, ввиду отсутствия связи с исследуемым показателем, из дальнейшего рассмотрения исключаются показатели «Количество новых

депозитов», «Средний чек в формате фастфуд», «Средний чек в формате ресторан», «Средние траты в ресторане».

Третьим этапом работы стал анализ матрицы коэффициентов парной корреляции с целью выявления мультиколлинеарности. В случае, если коэффициент корреляции между двумя объясняющими переменными превышает 0,8, одна из этих переменных, имеющая менее тесную связь с зависимой переменной, исключается из дальнейшего рассмотрения. В результате анализа матрицы коэффициентов парной корреляции были исключены из рассмотрения семь показателей: «Средняя сумма заявки на потребительский кредит», «Средняя сумма заявки на ипотечный кредит», «Средняя сумма нового депозита», «Средняя пенсия», «В среднем депозитов в руб. на человека», «Средние расходы по картам», «Средние траты в ресторане фастфуд» [2].

Таким образом, были выявлены показатели, имеющие наиболее тесную связь с количеством заявок на потребительские кредиты в Костромской области: «Количество заявок на ипотечные кредиты», «Средняя зарплата», «В среднем руб. на текущем счете на человека».

Были рассмотрены значения коэффициентов парной корреляции объясняющих переменных с зависимой и сделаны предположения о типе связи между ними [5].

Значение коэффициента корреляции для показателей «Количество заявок на ипотечные кредиты» и «Количество заявок на потребительские кредиты»:

$$r_{x_2y} = 0,637.$$

В виду того факта, что динамика изменения количества заявок на кредиты в целом развивается схожим образом, можно сделать предположение о том, что достаточно тесная связь между показателями обусловлена их зависимостью от иного фактора, например, от средней зарплаты.

Значение коэффициента корреляции для показателей «Средняя зарплата» и «Количество заявок на потребительские кредиты»:

$$r_{x_6y} = 0,484$$

Исходя из функций, выполняемых потребительским кредитом, можно сделать предположение о том, что связь между рассматриваемыми показателями обусловлена значительным влиянием среднего уровня заработной платы человека на его финансовое состояние. С одной стороны, в затруднительном финансовом положении человек прибегает к использованию потребительского кредита как к инструменту как к возможности справиться с текущими сложностями, с другой стороны, с повышением заработной платы человек начинает рассматривать различные способы улучшения уровня собственной жизни и использует потребительский кредит как инструмент для приобретения дорогостоящих товаров. Возможно, именно поэтому уровень связи между средней зарплатой и количеством заявок на потребительские кредиты составляет среднее значение, а не является линейной зависимостью.

Значение коэффициента корреляции для показателей «В среднем руб. на текущем счете на человека» и «Количество заявок на потребительские кредиты»:

$$r_{x8y} = 0,571$$

В ситуации с количеством денежных средств на счете человека, можно сделать предположение о том, что чем меньше денежных средств на счете человека, тем выше вероятность того, что он оставит заявку на потребительские кредит с целью преодоления временных трудностей. Степень связи между рассматриваемыми показателями достаточно высокая, что может быть обусловлено отсутствием у людей постоянного источника доходов. Если в случае с заработной платой, человек может спланировать свои расходы таким образом, чтобы избежать использования потребительского кредита, то у человека без работы такая возможность зачастую отсутствует.

Таким образом, в исследовании представлен алгоритм проведения совместного анализа, который можно использовать как при изучении деятельности предприятий малого и среднего бизнеса, так и крупных организаций. Исследование может быть полезным компаниям и предпринимателям, разрабатывающим собственную систему бизнес-аналитики.

### **Библиографический список:**

1. Грекул В. И. Проектирование информационных систем. — М.: РГГУ, 2018. — 224 с.
2. Информационные системы и технологии в экономике и управлении: Учебник для академического бакалавриата / В. В. Трофимов [и др.]; под ред. В. В. Трофимова. — 4-е изд., перераб. и доп. — М.: Издательство Юрайт, 2017. — 542 с.
3. Межов И. С. Сравнительная характеристика программно-аппаратных решений для обработки больших данных на российском рынке // Скиф. Вопросы студенческой науки. — август 2019. — №8. — С. 178-181.
4. Орлова И. В., Половников В. А. Экономико-математические методы и модели: компьютерное моделирование: Учеб. пособие. — М.: Вузовский учебник, 2017. — 365 с.
5. Шваб Клаус Четвертая промышленная революция: перевод с английского / Клаус Шваб. — Москва: Эксмо, 2019. — 209 с.