

Ивин Вячеслав Вадимович

канд. экон. наук, доцент

Кригер Александра Борисовна

канд. физ.-мат. наук, доцент

Школа экономики и менеджмента

ФГАОУ ВО «Дальневосточный федеральный университет»

г. Владивосток, Приморский край

DOI 10.31483/r-112484

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВЛИЯНИЯ ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ УНИВЕРСИТЕТОВ НА УРОВЕНЬ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ В РЕГИОНАХ РОССИЙСКОЙ ФЕДЕРАЦИИ

Аннотация: в статье предпринята попытка выявить статистическую зависимость между уровнем развития разработки отечественного программного обеспечения и результатами работы региональных вузов по подготовке выпускников в области ИТ.

Ключевые слова: высшее образование, информационные технологии, разработка программного обеспечения, статистический анализ.

Введение.

Анализ текущих тенденций на рынке информационных технологий (далее – ИТ) позволяет выявить три основные проблемы.

1. *Необходимость импортозамещения в сфере ИТ* обусловлена высоким уровнем зависимости от зарубежных разработчиков программного и аппаратного обеспечения. Это может привести к росту рисков непрерывности бизнеса многих крупных российских организаций.

2. *Дефицит кадров в ИТ-сфере* наблюдается по всем основным специальностям: разработчики, тестировщики, аналитики, системные администраторы, а также на управленческие кадры. Это приводит к росту заработных плат и, соответственно, увеличению стоимости ИТ-услуг для заказчиков.

3. Рост наборов абитуриентов на ИТ-направления в университетах связан с дальнейшим развитием ИТ-рынка и цифровизацией отраслей экономики России. Драйверами роста в ближайшие несколько лет станут импортозамещение и господдержка.

Данные.

В качестве исходных данных для исследования авторы использовали информацию о выпускниках государственных учреждений высшего профессионального образования, данные о зарегистрированных правах на программное обеспечение в «Российском реестре ПО», а также сведения о численности населения в различных регионах.

Источники данных – статистика высшего образования (сайт Минобрнауки [2]), Мониторинг качества приёма в вузы [4], Реестр российского программного обеспечения Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации [1].

Для анализа были выбраны данные по двум округам РФ: Центрального федерального округа (ЦФО), включая г. Москва, и Дальневосточного федерального (ДФО). Выбраны данные с 2021 по 2023 г.

Проблемы данных.

1) очень большой объём данных: исходные массивы данных (до выполнения запросов и предварительной обработки) составляют более 97 мегабайт. И это только для двух обозначенных федеральных округов;

2) статистические данные о студентах и выпускниках высших учебных заведений не подходят для статистического анализа без предварительной тщательной многоуровневой обработки и систематизации.

3) данные из реестра программного обеспечения представлены в виде нарастающего итога, что требует выполнения нескольких запросов.

Как следствие для анализа выбраны два округа: ЦФО – с наибольшим числом зарегистрированного ПО и ДФО («домашний» регион авторов) с достаточно скромными результатами в ИТ отрасли.

Цель исследования – выяснить, существует ли связь между количеством выпускников профильных специальностей в университетах и объёмом разработки программного обеспечения в регионах.

Описание датафреймов, сформированных для анализа.

Наборы данных формировались с помощью приложения *Loginot Community* исходя из сформированных гипотез для аналитики использовались три разных витрины данных (сформированных).

1. *Витрина со сквозными данными с 2022 по 2023 г.* (рис. 1). Включает данные по выпускникам ИТ-направлений подготовки (В этом исследовании авторами были рассмотрены выпускники, получившие образование в области ИТ по следующим направлениям подготовки: 09.03.01, 09.03.02, 09.03.03, 09.03.04, 02.03.01, 02.03.02, 02.03.03, 02.04.01, 02.04.02, 02.04.03, 09.04.01, 09.04.02, 09.04.03, 09.04.04, 01.03.02, 01.03.04, 01.04.02, 01.04.04, 38.03.05, 38.04.05) – государственные и не государственные образовательные учреждения, по указанным выше специальностям. Негосударственные учреждения не вносят значительного вклада в подготовку специалистов в этой области, разница в количестве выпускников составляет менее 1%.

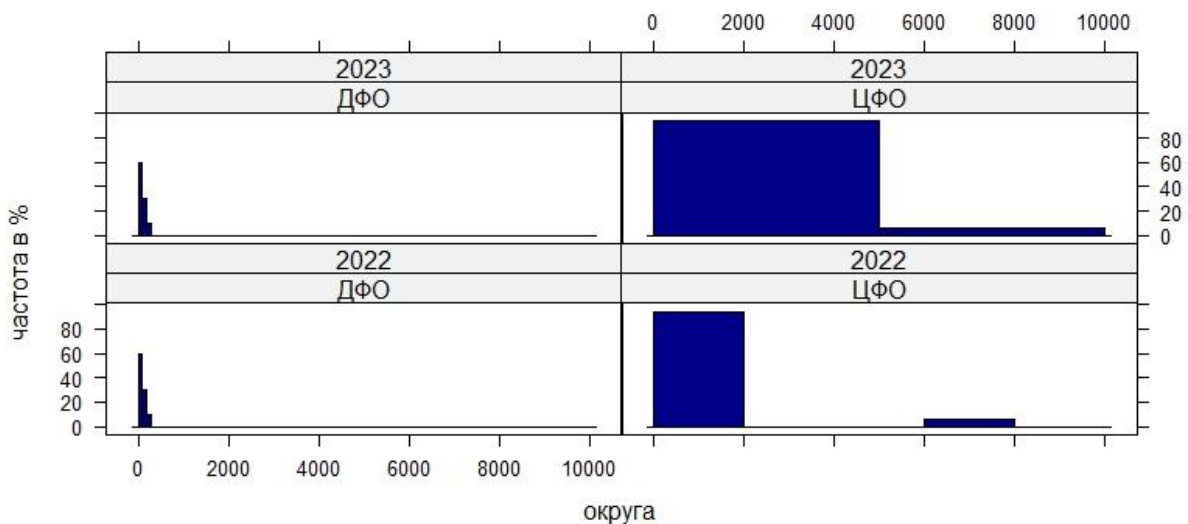


Рис. 1. Выпускник ИТ-направлений

2 Витрины данных отдельно для 2022 и 2023 гг. (рис. 2). Данные по выпускникам ИТ-направлений подготовки представлены отдельно за каждый год. В структуру данных добавлена численность населения региона, детализированные и укрупнённые группы специальностей;

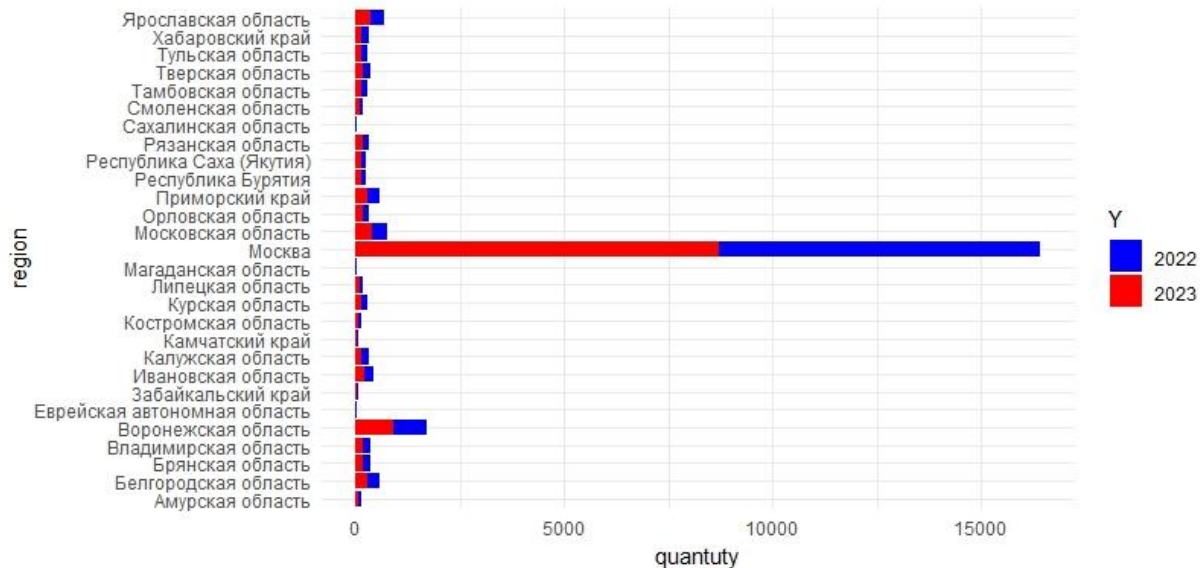


Рис. 2. Динамика изменения количества выпускников ИТ направления в регионах ЦФО и ДФО

3. Данные о зарегистрированных правах на программное обеспечение за 2021, 2022 и 2023 гг. (рис. 3). Здесь важно отметить, что большинство патентов на программное обеспечение было получено до конца 2021 г. Однако это не означает, что все упомянутые программы были выпущены именно в 2021 г. (рис. 4).

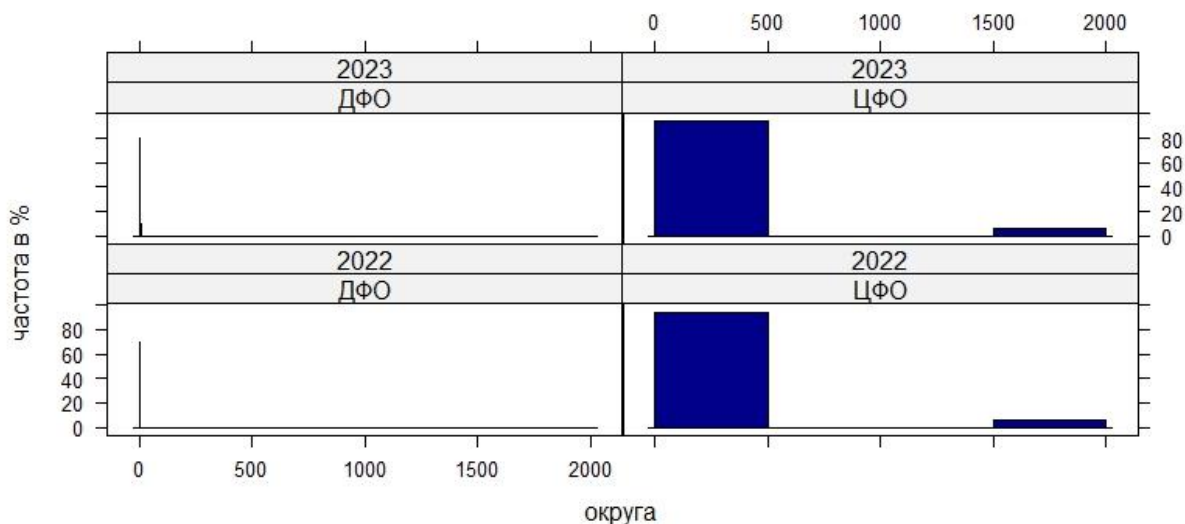
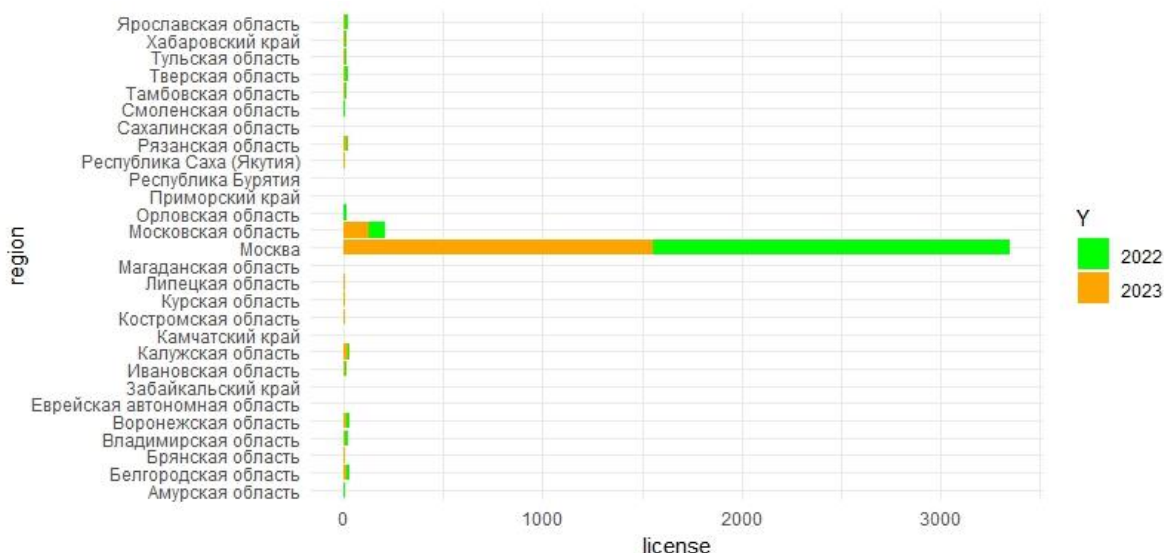


Рис. 3. ПО в реестре

Рис. 4. Динамика изменения зарегистрированного ПО
в регионах ЦФО и ДФО

Регистрация программного обеспечения в Реестре началась в 2017 г. Большинство регистраций было проведено в конце 2021 г., что связано с изменениями в законодательстве. Большая часть ПО, зарегистрированного до 2021 г., была создана ранее. Поэтому динамику новых разработок можно оценивать только на основе данных за 2022–2023 гг.

Учитывая ограниченное количество наблюдения по числу лет, полноценный анализ динамики показателей сделать не представляется возможным. Од-

нако, представленные диаграммы позволяют сделать вывод, что объём «выпуска» университетов в 2022 и 2023 гг. сохранились практически на одном уровне (по выбранным нами укрупненным направлениям подготовки). При этом, в исследовании [3] приведены статистические оценки численности студентов ИТ-специальностей в российских вузах (за 2015/2016, 2017/2018 и 2019/2020 учебные годы), согласно которых, контингент студентов существенно вырос. Авторы полагают, что такая разница в оценках связана с тем, что в указанной работе [3] использовались оценки по УГС 02.00.00 «Компьютерные и информационные науки», 09.00.00 «Информатика и вычислительная техника», 10.00.00 «Информационная безопасность» и 11.00.00 «Электроника, радиотехника и системы связи» в целом *по студентам, проходящим обучение*. Тогда как, авторы выделили только фактически выпущенных специалистов, по конкретному перечню направлений.

Количество зарегистрированного официально программного обеспечения меняется разнонаправленно по регионам. Очевидное лидерство Москвы, как по высшему образованию, так и по разработке программного обеспечения вполне естественно. Москва является финансовым центром РФ и крупнейшим городом не только страны, но и Европы.

В предварительном этапе исследования были сформулированы (выдвинуты) три базовые *гипотезы*.

1. *Гипотеза о том, что региональные компании используют внутренние кадровые ресурсы, т. е. выпускников университетов, расположенных на территории региона (и, возможно, сотрудников университетов).*

2. *Гипотеза о том, что основной кадровый состав находится в регионе регистрации компании, может быть связана с необходимостью обеспечения оперативности взаимодействия между разработчиками.*

3. *Резкий рост производства информационных компаний с 2022 года, связанный с уходом с российского рынка ИТ-разработчиков и США и Западной Европы.*

Статистический анализ по сквозной выборке.

Результаты корреляционного анализа для всей выборки представлены в таблице 1. Однако возникает закономерный вопрос: не влияет ли на результаты наличие данных по Москве в выборке? В дальнейшем будут представлены оценки, которые позволяют утверждать, что присутствие данных по Москве не влияет на обнаруженные закономерности.

Таблица 1

Результаты корреляционного анализа для всей выборки

	quantity	license	Y
quantity	1.000	0.9857	0.0147
license	0.986	1.0000	-0.0098
Y	0.015	-0.0098	1.0000

Корреляционная матрица подтверждает высокий уровень связи между количеством выпускников и числом разработанного, зарегистрированного ПО. Оценённая модель регрессии (показатели «федеральный округ» и «год» использовались как dummy-переменные) подтверждает влияние «объёма» выпуска университетов на количественную характеристику разработки ПО (результаты оценки представлены ниже). При этом объясняющие переменные «федеральный округ» и «год» оказались статистически не значимыми (таблица 2).

Таблица 2

Протокол регрессионного анализа для всей выборки

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	-8.623660	13.906065	-0.620	0.538
Quantity	0.204547	0.004829	42.358	<2e-16 ***
Y2023	-15.100732	14.265326	-1.059	0.295
StateЦФО	-14.113211	15.147638	-0.932	0.356
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 53.37 on 52 degrees of freedom				
Multiple R-squared: 0.9726, Adjusted R-squared: 0.971				
F-statistic: 614.7 on 3 and 52 DF, p-value: < 2.2e-16				

Полученные статистические оценки позволяют утверждать, что на уровень разработки ПО однозначно влияет результат образовательной деятельности университетов. Однако, возникает очевидная гипотеза, что количество выпускников является следствием численности населения в регионах.

Проверим очевидную гипотезу, что выявленная взаимосвязь является следствием численности населения в регионах, т. е. на самом деле данная взаимосвязь статистически не значима. Для этого учтём возможное влияние численности населения. Так как статистические данные по указанному показателю оценены на данный момент только за 2022 г., будем рассматривать отдельно массивы данных для 2022 и 2023 гг.

Для данных 2022 г. коэффициент частной корреляции (без учета влияния численности населения) между количеством выпускников и числом разработанного ПО составляет 0.995. Такой результат полностью опровергает гипотезу о влиянии «выброса». Соответственно для данных 2023 г. коэффициент частной корреляции (без учета влияния численности населения) между количеством выпускников и числом разработанного ПО составляет 0.994.

Таким образом, гипотеза о незначимости влияния количеством выпускников на объем, разработанного и зарегистрированного в регионах ПО очевидным образом отвергается.

Для подтверждения наших выводов были построены модели регрессии, объясняющей объём разработанного ПО, с учётом численности населения для данных 2022 (таблица 3) и 2023 гг. (таблица 4), соответственно.

Таблица 3

Протокол регрессионного анализа для 2022 г.

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	-20.816843	11.082867	-1.878	0.0725
Quantity	0.231439	0.008679	26.666	<2e-16 ***
StateЦФО	-20.698794	13.796554	-1.500	0.1466
Pop	0.002690	0.004706	0.572	0.5730
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 33.66 on 24 degrees of freedom				
Multiple R-squared: 0.9912, Adjusted R-squared: 0.9901				
F-statistic: 897.4 on 3 and 24 DF, p-value: < 2.2e-16				

Таблица 4

Протокол регрессионного анализа для 2023 г.

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	-20.610368	9.883755	-2.085	0.0479 *
Quantity	0.166331	0.006828	24.362	<2e-16 ***

StateЦФО	-17.407865	12.287862	-1.417	0.1694
Pop	0.010170	0.004193	2.426	0.0232 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 29.97 on 24 degrees of freedom				
Multiple R-squared: 0.9907, Adjusted R-squared: 0.9896				
F-statistic: 854.4 on 3 and 24 DF, p-value: < 2.2e-16				

Результаты оценки моделей регрессии ясно показывают, что переменная «численность населения в регионах» не имеет существенного влияния. Это позволяет сделать статистически подтверждённый вывод о том, что количество выпускников профильных направлений подготовки оказывает влияние на результаты работы компаний-разработчиков ПО.

Выводы.

Как для ЦФО, так и для ДФО уровень развития разработки программного обеспечения, выраженный в количестве зарегистрированного ПО, статистически значимо связан с результатами работы региональных университетов.

Авторы этой работы не имеют информации о кадровом составе разработчиков. Вероятно, предприятия не предоставляют такие сведения при регистрации своих разработок. Поэтому мы не можем с уверенностью сказать, что компании, разрабатывающие программное обеспечение, нанимают только сотрудников с профильным образованием. Однако наши результаты позволяют сделать вывод, что региональные университеты оказывают прямое или косвенное влияние на развитие разработки в ИТ-сфере.

Вопрос о том, какие направления подготовки в университетах действительно способствуют успешному импортозамещению программного обеспечения в России, остаётся открытым. Авторы планируют ответить на него в ходе следующего исследования.

Список литературы

1. Аналитический отчёт, отображающий количество правообладателей по субъектам РФ и количество ПО, содержащееся в реестре в период с 01.01.2024 по 02.07.2024 // Реестр российского программного обеспечения [Электронный

ресурс]. – Режим доступа: <https://reestr.digital.gov.ru/analytics/owners-region-map/> (дата обращения: 01.06.2024).

2. Высшее образование // Министерство науки и высшего образования Российской Федерации [Электронный ресурс]. – Режим доступа: <https://minobrnauki.gov.ru/action/stat/highed/> (дата обращения: 01.06.2024).

3. Меликян А.В. Подготовка IT-специалистов в российских вузах: статистический анализ / А.В. Меликян // Вопросы статистики. – 2022. – №29 (6). – С. 74–83. – DOI <https://doi.org/10.34023/2313-6383-2022-29-6-74-83>. EDN NTHYTW

4. Мониторинг качества приёма в вузы // Национальный исследовательский университет «Высшая школа экономики» [Электронный ресурс]. – Режим доступа: <https://ege.hse.ru/> (дата обращения: 01.06.2024).