



**ОДИНЦОВ
АЛЕКСАНДР ВЛАДИМИРОВИЧ**

кандидат социологических наук,
руководитель аналитического отдела издания
«Новости GXP»

«ЧЕРНЫЕ ЯЩИКИ» И «СЖАТЫЕ» ДАННЫЕ: ПОЧЕМУ РЕЙТИНГИ ТЕРЯЮТ ЗНАЧИМОСТЬ ДЛЯ БИЗНЕСА

(НА ПРИМЕРЕ РЕЙТИНГА СИСТЕМЫ ЗДРАВООХРАНЕНИЯ)

В деловых исследованиях рейтинги зачастую имеют «развлекательные» функции. Исключение составляют финансовые рейтинги, создаваемые мировыми агентствами. В настоящий момент рейтинги в здравоохранении и фармацевтической отрасли не являются существенными ориентирами для делового сообщества. Ниже в качестве одного из примеров будет рассмотрен рейтинг, сформированный в рамках «Исследования эффективности здравоохранения в городах мира», разработанный компанией «большой четверки» — PricewaterhouseCoopers (PwC)¹. Он выбран по двум основным причинам — относительная открытость методологии и заявленная низкая зависимость от экспертных методов. Исследование проводилось в 2018 году в 15 мегаполисах мира. Данными для формирования рейтинга послужили открытые статистические данные, отчеты глобальных аналитических агентств, результаты массового опроса, проведенного в городах исследования, и экспертные оценки.

Все ключевые методологические вопросы к формированию и использованию рейтингов можно объединить в несколько групп:

- 1) отбор критериев для расчета рейтинга и параметров, по которым эти критерии формируются;
- 2) отбор объектов формирования рейтинга;

¹ Исследование эффективности здравоохранения в городах мира // PwC. 2018. URL: <https://www.pwc.ru/ru/publications/health-research.html>.

- 3) использование экспертных оценок (насколько от них зависит конечный результат и как формируется пул экспертов);
- 4) процедуры сбора условно объективных, статистических данных и данных социологических исследований;
- 5) непосредственно расчет рейтинга и его интерпретация;
- 6) объемы «сжатия» данных (мера потери информации при проведении процедур анализа).

Рассмотрим их непосредственно в применении к исследованию РwС.

1. Отбор критериев для расчета рейтинга и параметров, по которым эти критерии формируются.

Очевидно, что именно от них зависит конечный результат, а непосредственно их количество и возможность их объективной оценки еще более значимы при рассмотрении сложных макросоциальных объектов (в нашем случае — мегаполисов мира). Аналитики РwС выбрали пять факторов, выделенных Всемирной организацией здравоохранения (ВОЗ) и определяющих здоровье нации². Но в методологии ВОЗ эти факторы являются скорее теоретической рекомендацией, а не жесткой установкой. Аналитики РwС не только выбрали для трех интересующих их критериев конкретные описывающие их параметры, но и укрупнили их до двух (организация медицинской и немедицинские факторы).

Также аналитики добавили еще один критерий — «индекс здоровья жителей», в который вошли коэффициент смертности в трудоспособном возрасте, ожидаемая продолжительность жизни при рождении, доля детей, родившихся с низкой массой тела. В методологических рекомендациях ВОЗ и материалах, на которые ссылается организация, параметров изучаемой сферы много больше, а сопоставление некоторых выбранных параметров достаточно спорно. Например, в Москве в силу постсоветской специфики организации медицинских учреждений длительность пребывания в стационаре заметно больше, чем в Стамбуле или Нью-Йорке, в которых из-за длительной тенденции на сокращение затрат на содержание койко-места пребывания в стационаре сокращалось год от года. Такой параметр, как количество койко-мест в медицинских учреждениях, вообще не вошел в рейтинг, как и, к примеру, время ожидания прибытия скорой помощи и еще множество параметров оценки здравоохранения.

Точно так же и с немедицинскими факторами, среди которых аналитики РwС выбрали достаточно ограниченный набор критериев оценки. По итогам рейтинг ссылается на рекомендации ВОЗ, но по факту ведет себя в их отношении до-

²Там же. С. 15.

статочно волонтаристски, произвольно ограничивая их набор. Следовательно, отбор критериев проводился экспертами РwС по некой закрытой методологии.

2. Отбор объектов формирования рейтинга.

При отборе городов для участия в рейтинге было проведено несколько последовательных и достаточно обоснованных процедур, но «формирование окончательного списка проводилось с участием экспертов по здравоохранению, на основе результатов обсуждения с ними сопоставимости профилей городов по демографическим и инфраструктурным параметрам, характеризующим как города в целом, так и их системы здравоохранения»³. Города отбирались в рамках пяти последовательных этапов.

Если первые четыре этапа в той или иной мере можно считать объективными, то завершающая процедура полностью экспертная. Предварительные процедуры фактически были только подготовкой материалов для экспертных оценок. Это не значит, что объективность экспертов сразу ставится под сомнение. Но очевидно существует водораздел между, например, предварительными условиями для отбора (списком стран ОЭСР или количеством жителей города), которые может увидеть каждый и которые не зависят от РwС, и данными, полученными на неизвестной выборке экспертов при применении неизвестных методов обработки оценок.

3. Использование экспертных оценок (насколько от них зависит конечный результат и как формируется пул экспертов).

Само использование экспертных процедур оценки подвергает сомнению результативность рейтинга, так как автоматически вызывает вопросы к формированию выборки экспертов. Кстати, при проведении «Исследования эффективности здравоохранения в городах мира» аналитики РwС критиковали прочие «медицинские» рейтинги в основном именно за чрезмерное доверие мнению экспертов. Но и они активно прибегают к экспертному знанию, а методики формирования пула экспертов и обработки полученных от них мнений не раскрыты.

4. Процедуры сбора условно объективных, статистических данных и данных социологических исследований.

При формировании рейтинга эксперты РwС указывали, что доступной была лишь часть данных о городах исследования. Мало того, именно отсутствие значимого количества данных вынудило экспертов исключить из рассмотрения часть городов. Социологическая часть сбора данных, как правило, остается в тени. В рассматриваемом нами исследовании методология опроса населения в круп-

³ Там же. С. 16.

нейших городах мира не прописана, то есть нет никакой возможности оценить ни процедуры формирования выборки, ни влияние возможных выбросов данных на конечный результат опроса.

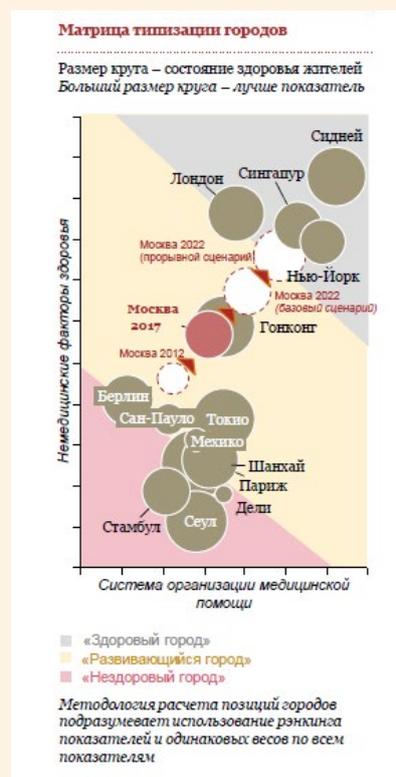
5. Непосредственно расчет рейтинга и его интерпретация.

По каждому критерию исследования города ранжировались по рассматриваемым показателям (худший — 1 балл, лучший — x баллов, где x — количество городов). Затем общий балл города по параметру считался путем вычисления среднего балла в каждом отдельном ранге по каждому из показателей, входящих в него. Кроме получения количественных показателей по отдельным критериям, аналитики PwC применили оригинальное решение — итоговые данные рейтинга были представлены на трехмерном графике («матрице»), в котором по оси Y учитывался индекс в сфере организации медицинской помощи, а по оси X — индекс немедицинских факторов здоровья. Третье измерение было представлено в виде размеров кружка, в котором нашел отражение еще один индекс — индекс здоровья жителей (чем больше кружок, тем выше значение индекса).

Результат выглядит следующим образом⁴.

Как видно из приведенной иллюстрации, фактически исследование PwC работает с тремя рейтингами: организации медицинской помощи, немедицинских факторов общественного здоровья и индекса здоровья граждан. Итоговый результирующий рейтинг существует в ограниченной порядковой шкале из трех делений — «здоровый город», «развивающийся город», «нездоровый город».

На основании сформированного рейтинга городов аналитики PwC находят факторы, которые влияют на то, является ли система здравоохранения «эффективной или нет». Интерпретация данных — дело качественное (используя принятую в отечественной социологии дихотомию «количественное» VS «качественное»). Здесь интерпретация рейтинга служит скорее вспомогательным инструментом для описания другого, построенного на экспертных методах исследования — «Ключевые тренды» в здравоохранении⁵.



⁴ Там. С. 30.

⁵ Там же. С. 20–27.

б. Объемы «сжатия» данных (насколько часто одни количественные данные заменяются другими, как используется математическая обработка экспертных методов, применяются ли веса или другие коэффициенты при обработке количественных данных).

В рассматриваемом нами исследовании наиболее интересны с этой точки зрения переходы от и без того сжатых данных официальной статистики к ранжированным рядам объектов, то есть к шкале, которая выглядит как числовая и порядковая, но игнорирующая существенные различия между объектами, заменяя их рангами. То есть если в исходных данных аналитики РwС имеют дело с интервальными шкалами, которые позволяют понять не только как объекты анализа расположены по размеру исследуемого признака, но и насколько они различаются, то для расчета каждого субрейтинга (критерия) рассматривается их ранжированный ряд — по определению по шкале более «низкого типа», которая позволяет понять только взаимное расположение объектов. Также следует помнить, что значение каждого критерия оценивается как среднее рангов, то есть авторы еще раз сжимают данные, сокращая количество значений, которыми оперируют.

Но и на этом «сжатие» не заканчивается, ведь итоговый рейтинг сводится к еще одному упрощению шкалы — делению всего списка на три типа городов: «здоровый город», «развивающийся город», «нездоровый город».

Исследование РwС — масштабный проект, методология которого значительно раскрыта, показанные выше проблемные моменты его формирования и интерпретации отлично иллюстрируют «родовые пятна» большей части рейтингов в целом. В сущности, ключевых проблем две:

1. Использование экспертных оценок.

Нередко рейтинги оперируют числовыми значениями, но они получаются благодаря использованию экспертных оценок. Использование чисел при обобщении экспертных мнений является частой практикой при обработке результатов экспертных опросов, но именно оно формирует иллюзию «объективности» полученных рейтингов. Это позволяет искусственно снимать вопросы о формировании пула экспертов исследования. Именно на таких непрозрачных методах получения значений строятся многие политические и околополитические рейтинги (например, индекс восприятия коррупции от Transparency International или индекс демократии от Economist Intelligence Unit). Экспертные оценки в рейтингах — это «черный ящик», генерирующий ответы по методологии, которая заявляется как понятная, но таковой не является.

В анализируемом исследовании РwС непосредственно для расчета значений экспертные оценки не использовались — по крайней мере, об этом не заявлено.

Но они активно применялись при формировании критериев и их параметров, отборе городов для исследования и, самое главное, при интерпретации данных исследования, когда становится очевидно, что проведенные процедуры по получению количественных значений показателей несущественны. Все проведенные расчеты должны быть только иллюстрацией экспертизы РwС в области здравоохранения. Здесь надо принять во внимание, что это исследование действительно имело PR-характер. Один из участников «большой четверки» получает доходы от других исследований, а открытые рейтинги служат привлечению внимания потенциальных клиентов.

2. Чрезмерность сжатия данных.

Всякий анализ данных, как и всякая процедура измерения, является процессом «сжатия данных», при которой для интерпретации информации мы сознательно убираем из рассмотрения отдельные стороны изучаемого явления. В отдельных случаях потеря информации для исследования может показаться приемлемой — как, к примеру, игнорирование в социально-демографических блоках таких характеристик, как рост респондента или его вес. В другом случае потеря информации может быть критической.

Ранжирование, основа любого рейтинга — это сжатие данных. Очевидно, что в основе всякого ранжирования и, следовательно, рейтингования стоит некая количественная шкала, которая позволяет выстроить объекты в определенном порядке, как правило, по убыванию изучаемого свойства. При этом в итоговом рейтинге само значение исследуемого признака не является чем-то важным, а становится факультативным. Результат рейтинга — новая порядковая шкала, на которой выстроены объекты наблюдения. Упрощая: неважно, какое значение индекса у Москвы, важно, какое место она занимает в общем ряду — кто выше нее, а кто ниже.

Потому всякий рейтинг — это попытка оценить сложное явление одной цифрой. Сблазн максимально упростить ситуацию, чтобы получить простые ответы на простые вопросы: «так всё плохо или хорошо?» и «а как у соседа?». В рассматриваемом нами рейтинге процедура сжатия данных проходила несколько раз. Сжатые статистические данные (в том числе результаты опроса) были сжаты в ранжированный ряд для расчета критерия, который после прошел сжатие с помощью расчета среднего значения, чтобы затем быть помещенным на трехмерную «матрицу», которая все равно потом оказалась отраженной на простой трехчленной порядковой шкале (плохо — нейтрально — хорошо, а в нашем случае — «нездоровый город», «развивающийся город», «здоровый город»).

Рейтинги в системе здравоохранения, как и многие прочие рейтинги, не могут быть основанием для сложных KPI крупных бизнес-организаций (сетевых кли-

ник, фармацевтических производителей и дистрибьюторов, производителей медицинского оборудования и т. д.), то есть служить ориентиром в практической деятельности, так как:

1. Процедуры получения рейтингов слишком привязаны к «черному ящику» экспертных методов, то есть проверить валидность этих методик и тем более самостоятельно провести какое-то исследование по аналогии с ними практически невозможно.
2. Сжатие данных в рейтинге чрезмерно — один параметр (собственно значение рейтинга) невозможно использовать для постановки задач всем подразделениям корпорации. Напротив, в бизнесе востребованы более сложные модели, например регрессионные, в которых конечный результат (итоговый KPI-прибыль) непосредственно математически связан с другими результатами деятельности (узнаваемостью продукта, медиаактивностью, эффективностью R&D, производительностью труда и т. д.).

Рейтинги — попытка слишком просто ответить на вопрос о динамике сложного явления. Именно эта чрезмерная простота ответа и оставляет им только развлекательные функции в сфере бизнес-исследований.

