

Федеральное государственное образовательное бюджетное учреждение
высшего профессионального образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)

Департамент анализа данных, принятия решений
и финансовых технологий

Дисциплина «Программирование в среде R»

П.Б. Лукьянов

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ
КОНТРОЛЬНОЙ РАБОТЫ № 1

Расчет показателей эффективности торговой сети
Часть 3

Для студентов, обучающихся по направлению подготовки
«Прикладная информатика»
(программа подготовки бакалавра)

Москва 2020

Генерация исходных данных средствами R

После выполнения предыдущих шагов были получены расчетные значения, результаты представлены в таблице. По этим показателям можно провести некоторый анализ: сравнить работу магазинов, выяснить удачные и убыточные дни и т.д. Казалось бы, работа закончена. Вместе с тем, у руководителя, аналитика возникают новые вопросы, например:

- Как изменятся показатели, если после проведения рекламной кампании продажи возрастут на 30, 50, 100 процентов?
- Какие цифры будут получены, если за час до закрытия не проданный товар будет продаваться со скидкой в 20, 30, 50 процентов?
- Что покажут расчеты, если одновременно с первым товаром продавать аналогичный, но из другой ценовой категории?

Таких вопросов, отражающих задачи развития бизнеса, текущую ситуацию на рынке, влияние сезонных факторов, может быть очень много. Для того, чтобы ответить на эти вопросы, оценить возможный результат, необходимы новые, специально подготовленные исходные данные – данные, имитирующие разные условия, различные сценарии и ситуации.

Чтобы создавать такие данные и получать различные прогнозы, используется имитационное моделирование – метод исследования, заключающийся в том, что реальные данные заменяются специально сконструированными, с достаточной точностью описывающими моделируемую ситуацию.

Поэтому для аналитика отдельным и важным навыком является умение получать «правильные», модельные данные, соответствующие возможному или запланированному варианту развития событий. Численные эксперименты и расчеты с модельными данными называют имитацией, поскольку реальные данные не используются.

В начале выполнения работы было сказано, что исходные данные по поставкам и продажам предоставляются некоторой информационной системой и отражают работу магазинов за неделю. Теперь, используя язык R, нужно научиться самостоятельно создавать любые наборы данных под заданные условия.

Задания

1. Написать программу по генерации файлов исходных данных (in.txt, out.txt) для всех магазинов. Генерировать данные, используя функции `runif()` и `sample()`. Для каждого магазина задать свой интервал возможных значений поставки товара. Проверять условие, чтобы объем продаж магазина не превышал объем поставок товара. Если обнаружена ошибка в исходных данных, менять данные местами или уменьшать продажи в соответствии с величиной поставки.
2. При генерации данных не использовать параметры, задающие точные границы случайных величин. Использовать решение задач 6 и 7 из Методички № 4.
3. Для каждого магазина создать константу (параметр) `saleLevel` с диапазоном возможных значений от 0 до 100. Данная константа определяет успешность работы магазина. Так, при значении `saleLevel = 0` разность между поставкой и продажей товара максимальна (товар вообще не продается), при `saleLevel = 100` весь полученный товар продается. Для разных магазинов задать произвольные значения `saleLevel`, например, 21, 77, 16, 54 и т.д. Выяснить вид связи `saleLevel` и прибыли магазина, `saleLevel` и рентабельности.
4. В файлах исходных данных создавать заголовки из двух слов, вместо «День» создавать заголовок «День недели», проверить

корректное считывание заголовков в основной программе обработки и анализа данных.

5. В файлах исходных данных вместо номеров дней недели создавать столбцы «понедельник», «вторник» и т.д. Соответствующим образом изменить основную программу обработки данных для работы с такими файлами.