

БЫТЬ ПОНЯТНЫМ

Андрей Гаськов

По мотивам «Storytelling with Data», Cole Knaflic

«Наши data scientist'ы понарисовали графиков,
давайте их куда-нибудь пристроим»

«Наши data scientist'ы превращают данные в
информацию для принятия решений»

Контекст

Визуализация

Мусор

Фокусировка

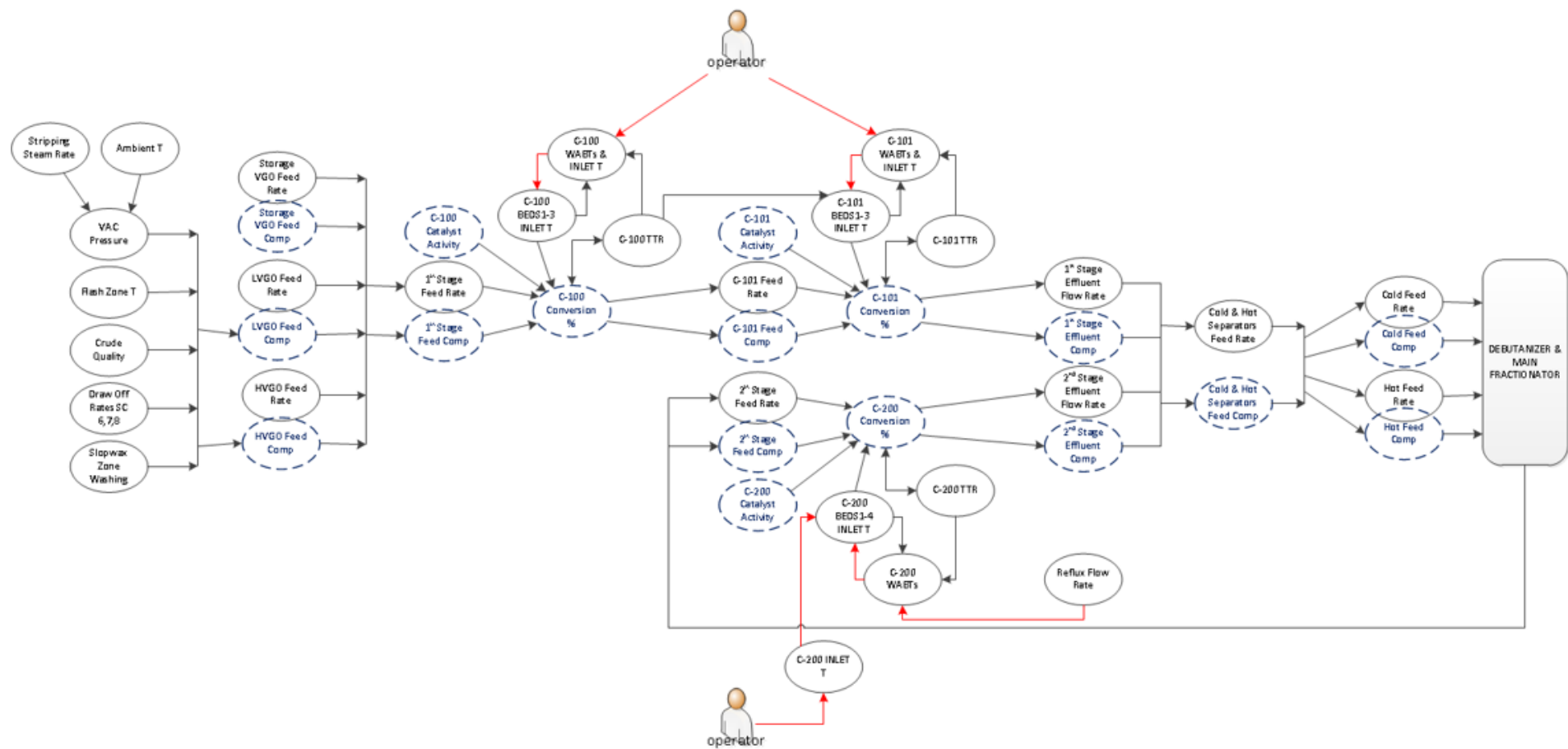
Дизайн

История

Контекст

Исследовательский анализ
(Exploratory Analysis)

Объясняющий анализ
(Explanatory Analysis)



Кто

Аудитория

Вы

Что

Необходимо знать

Призыв к действию

Как

Данные

Визуализация

эффективные способы

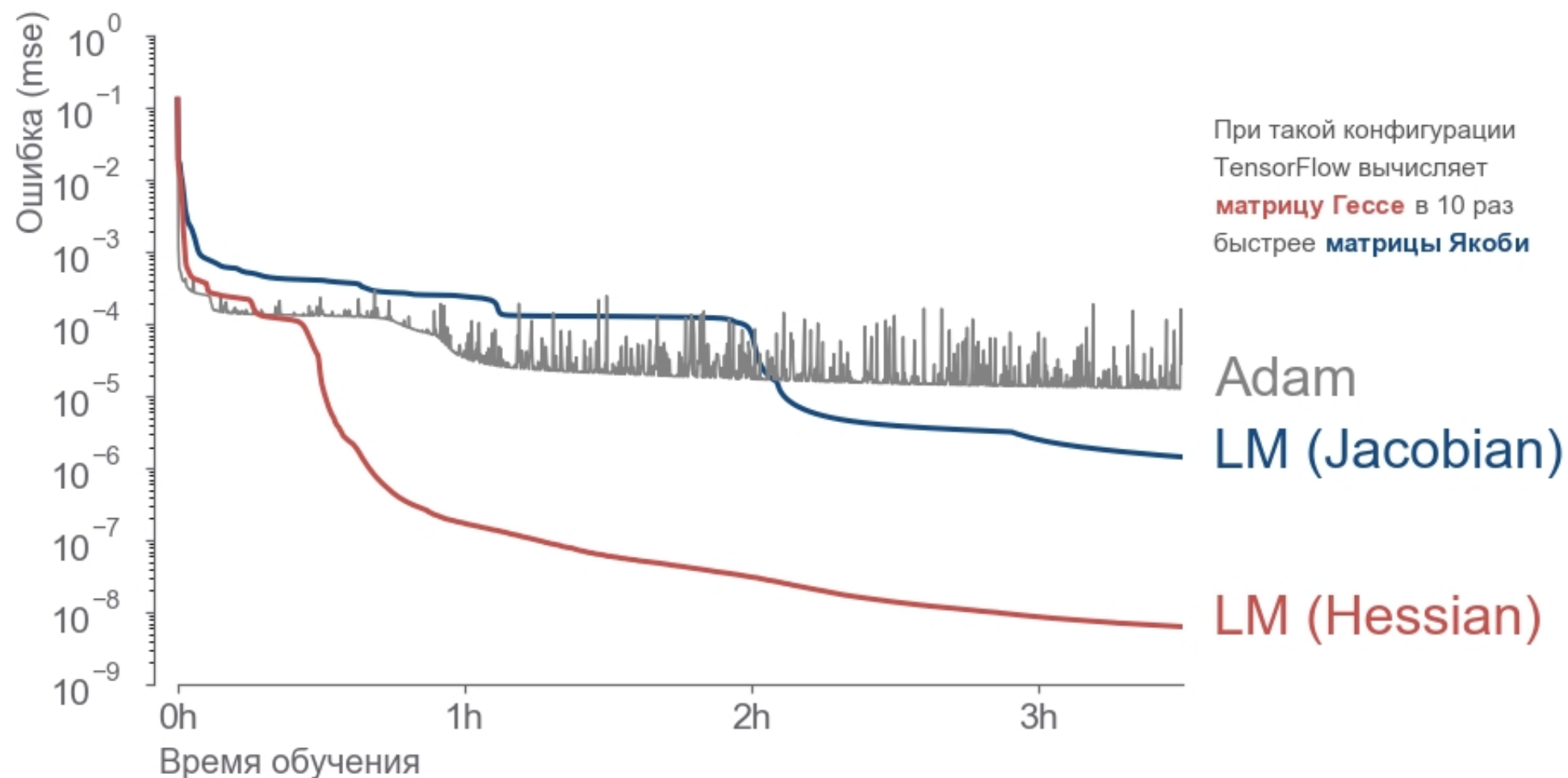
23%

студентов, поступивших в 1998-2001гг,
было отчислено из университета

По данным статьи Е. Донец "Опыт исследования студенческих отчислений на примере МГУ", 2011. 7,224 студента 1998–2001 гг. поступления.

	Юноши	Девушки
1-й курс	8%	8%
2-й курс	8%	3%
3-й курс	2%	3%
4-й курс	3%	2%
5-й курс	2%	2%

Алгоритм Левенберга-Марквардта (LM) vs Adam



Данные: 10K точек 6D функции Розенброка. Архитектура: MLP, 3 скрытых уровня, ширина 12, 10, 8 (311 параметров).
Количество шагов: Adam 2,445,187; LM (Jacobian) 1,468; LM (Hessian) 13,396.

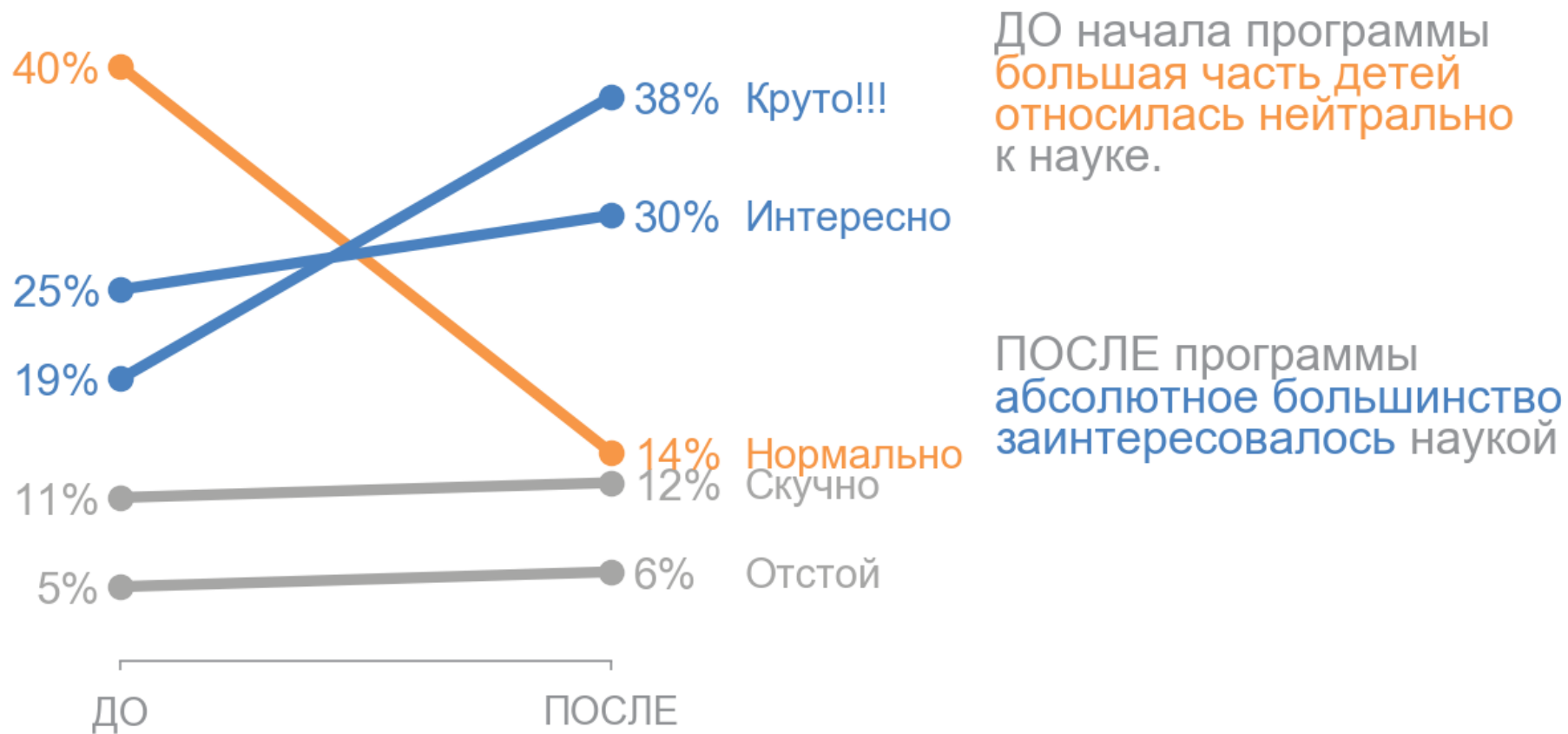
Стоимость под Риском vs. Ожидаемая Прибыль

за 1 день на \$10,000 инвестированных



Программа знакомства детей с наукой успешна

Опрос: "Наука - это ..."



На основе опроса NN детей в возрасте XX-YY лет, участвовавших в программе (100% ответивших в обоих опросах).

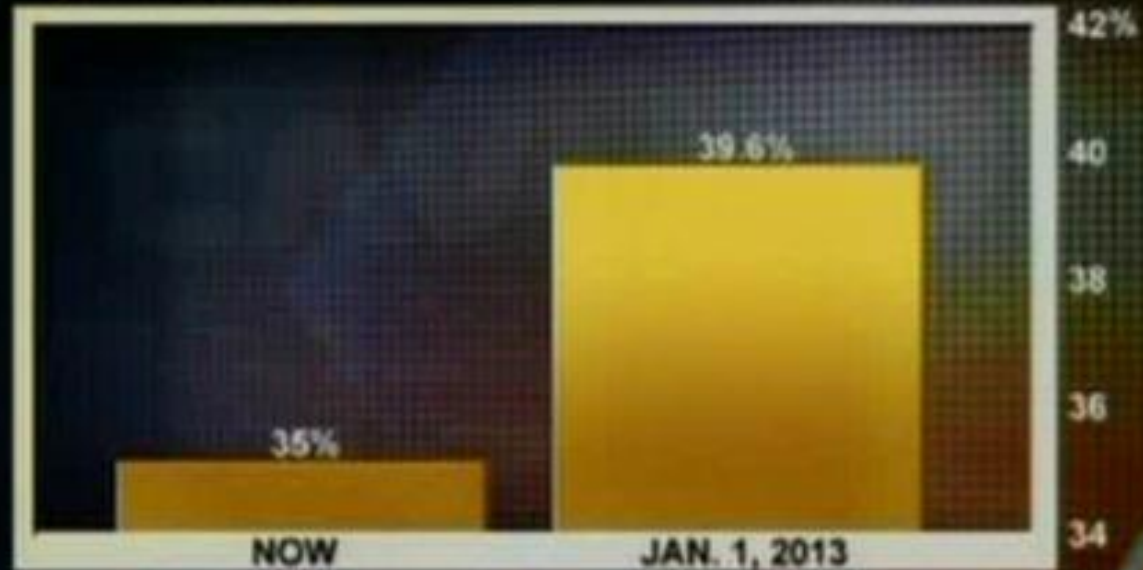
Если истечёт "бушевское" снижение налогов

максимальная ставка налога



IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



8:01p ET

FOX
BUSINESS

TOP STORIES

TECHNOLOGY

CONSUMER

WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

DOW 13008.68 ▼ 64.33

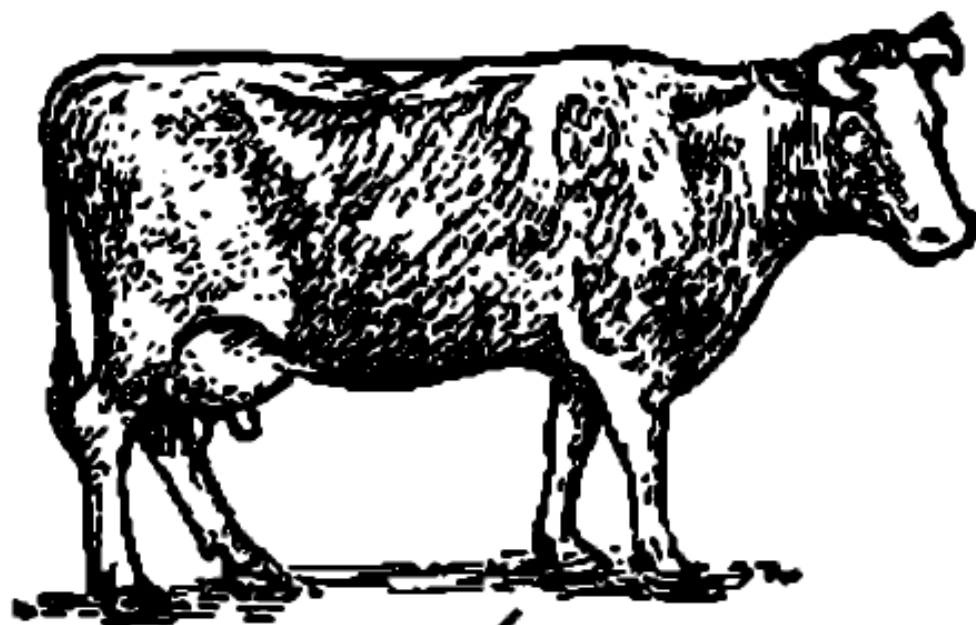
S&P 1379.32 ▼ 5.98

NASDAQ 2939.52 ▼ 6.32

THE CRESCIVE COW



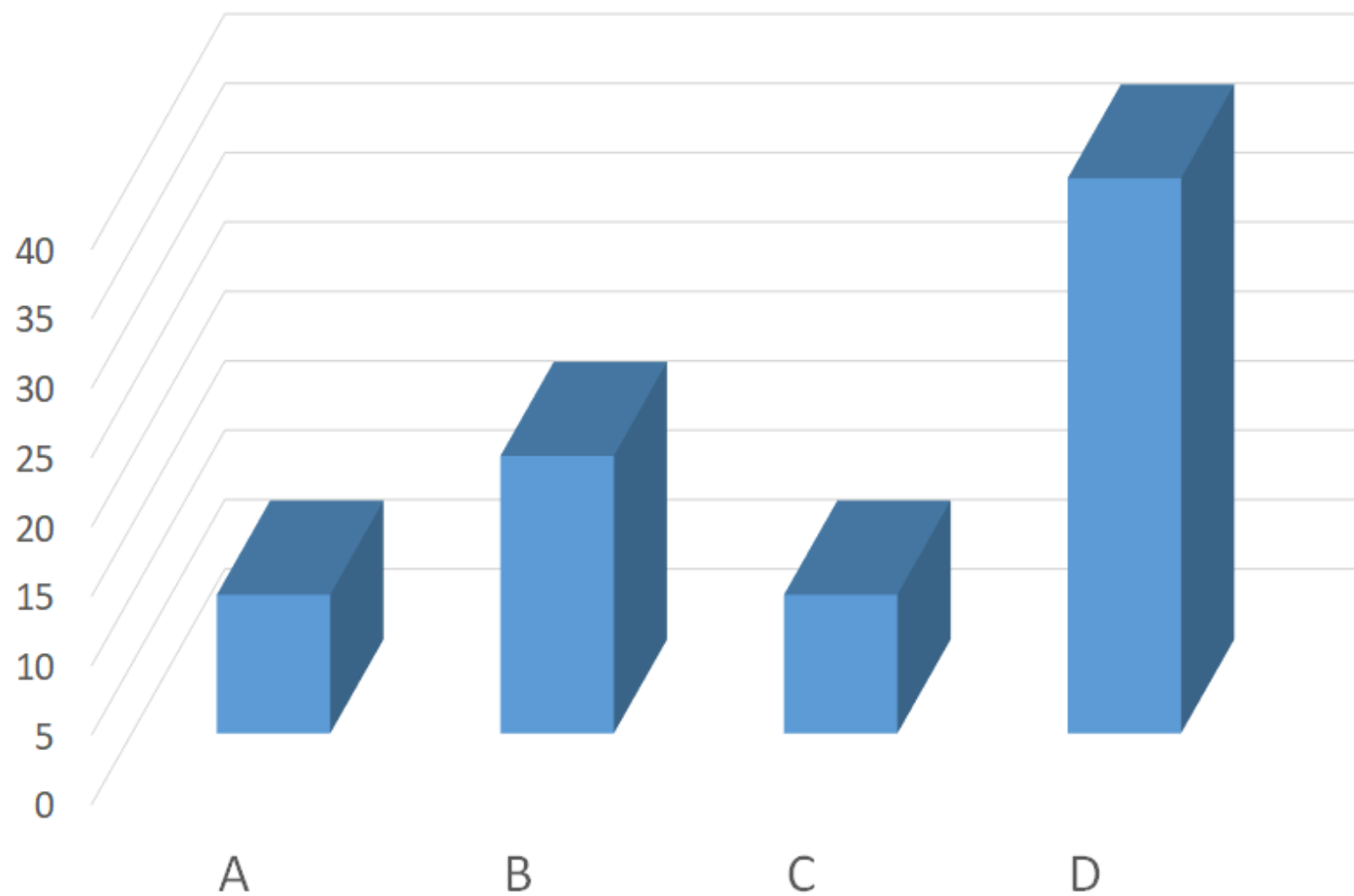
1860



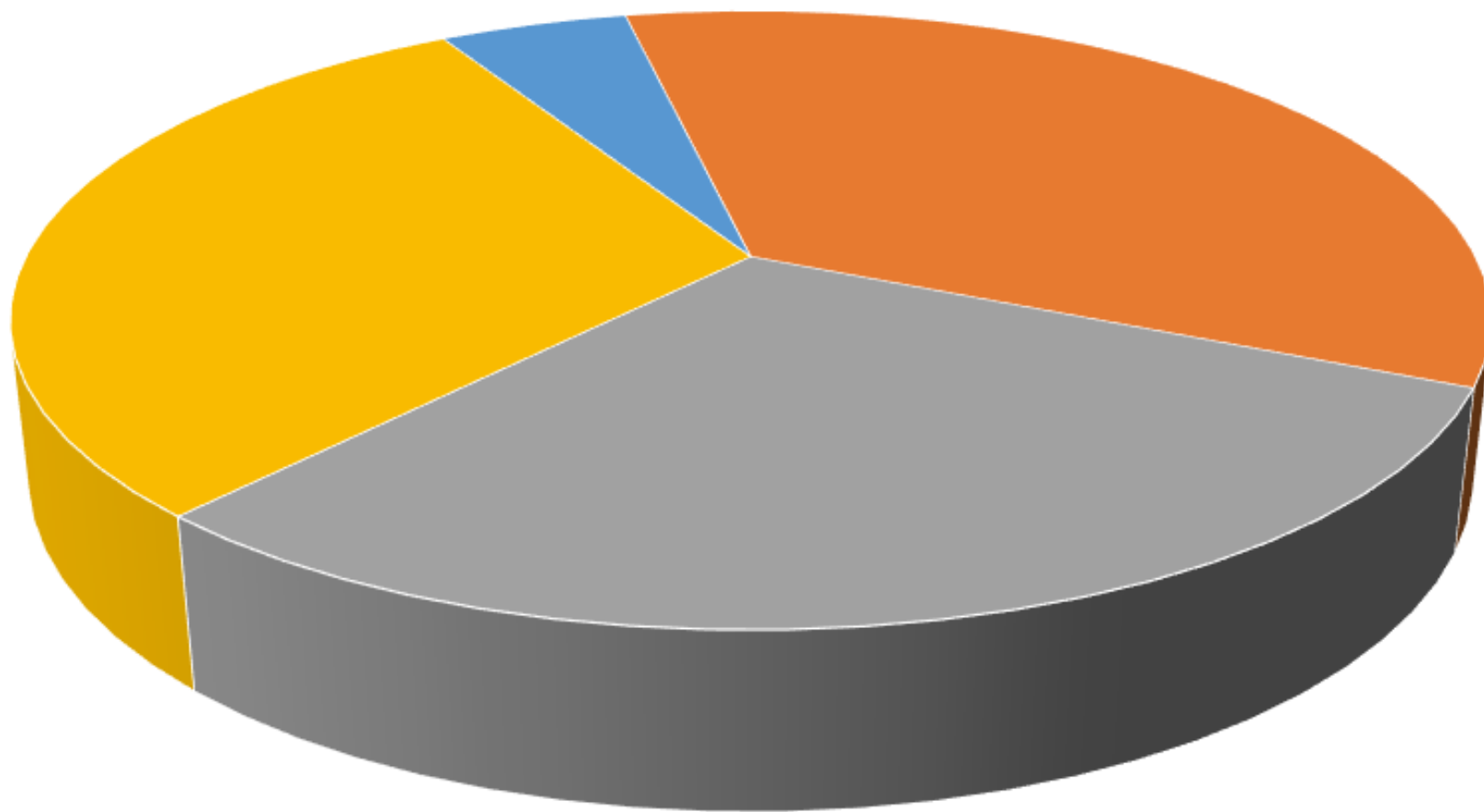
1936

Визуализация
избегать

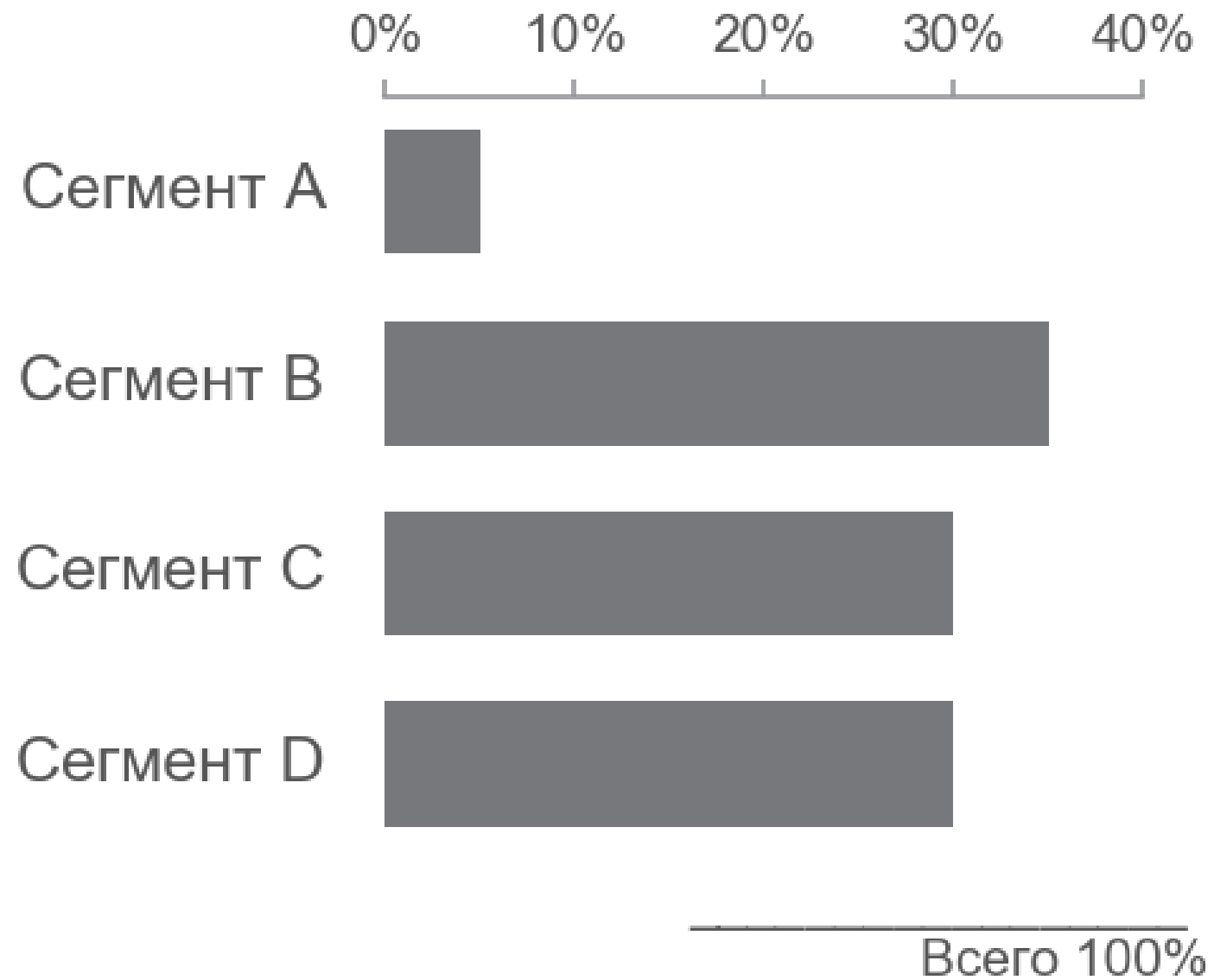
Какая высота столбца D?



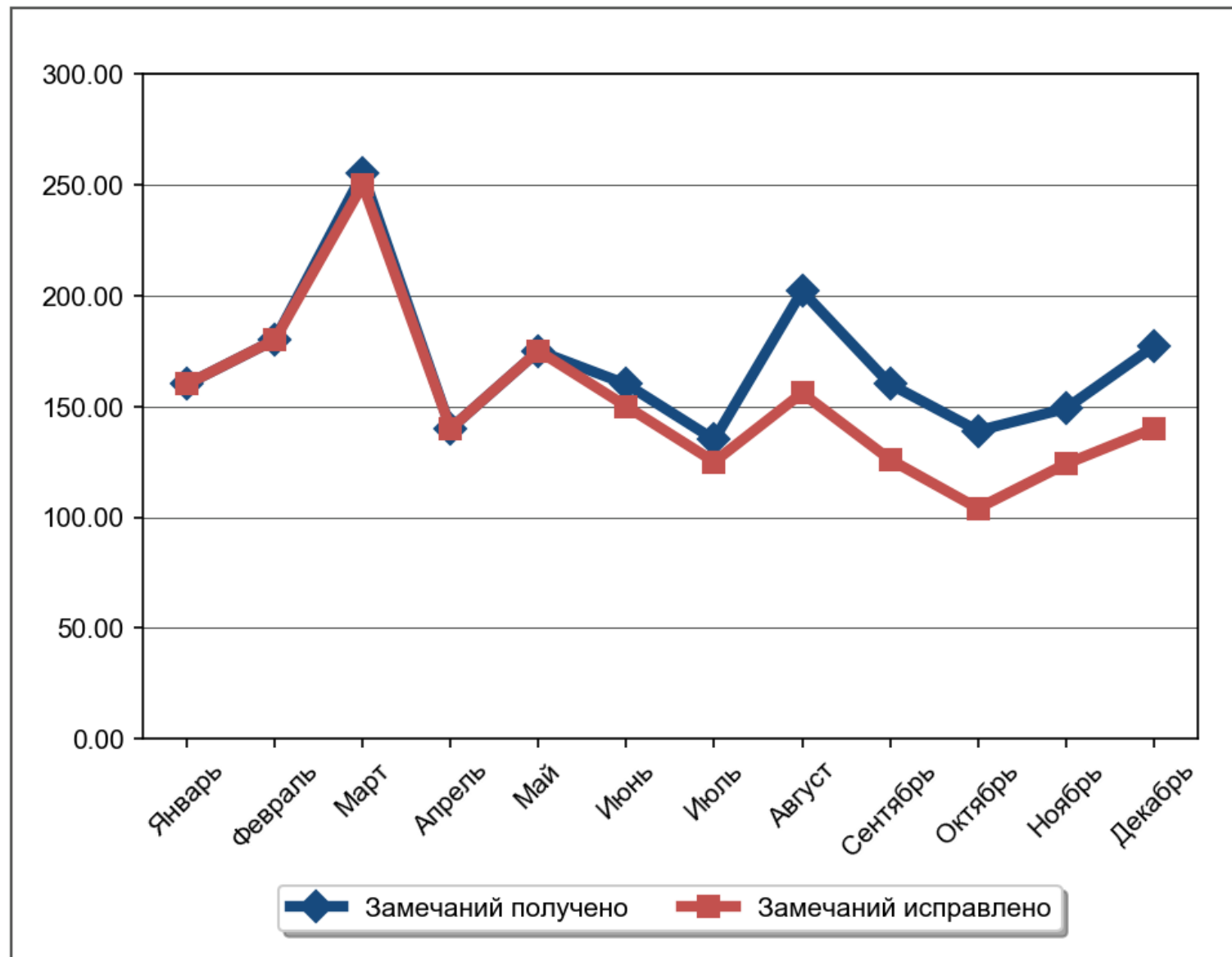
Какой сегмент больше, В или С?

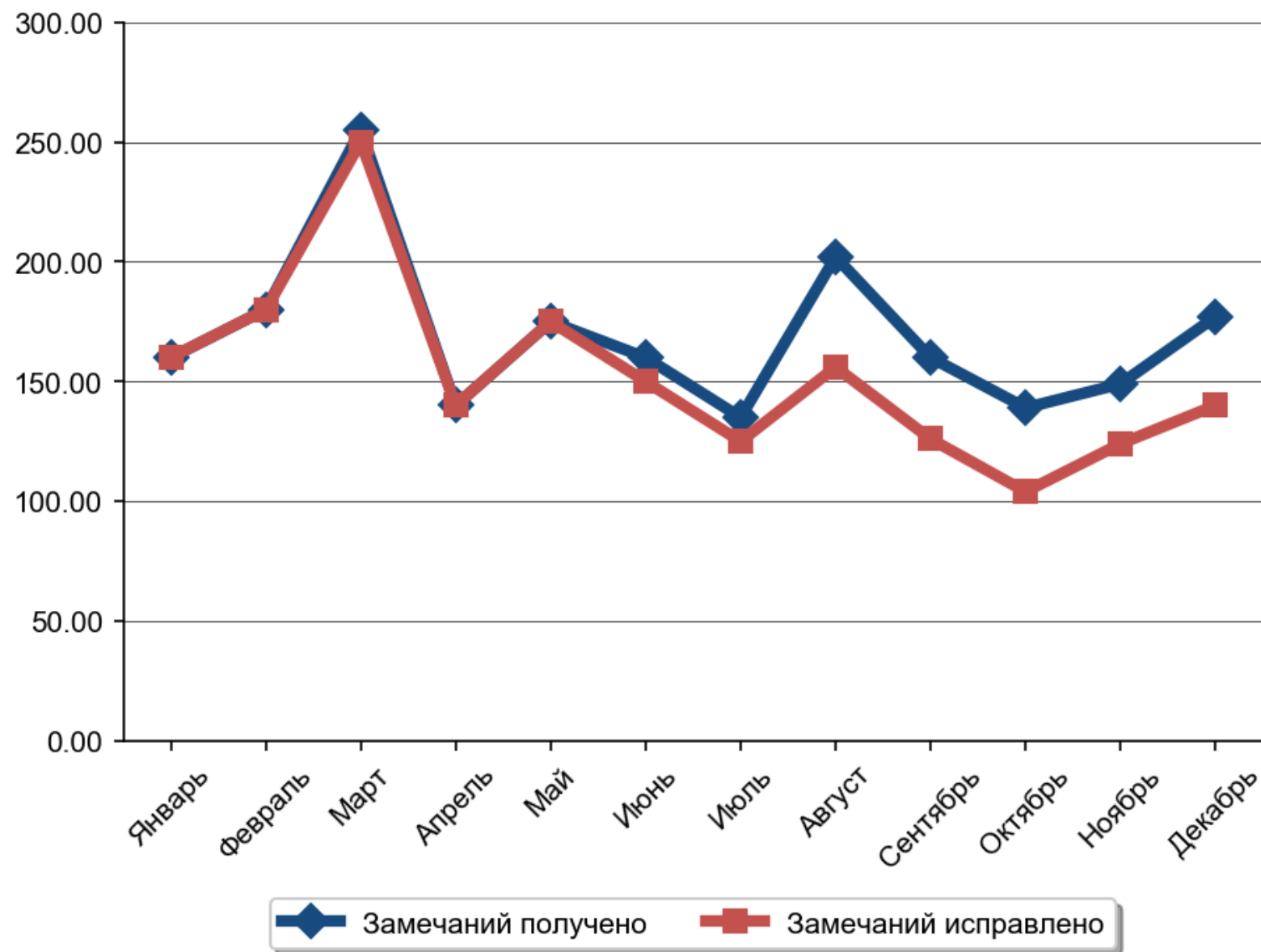


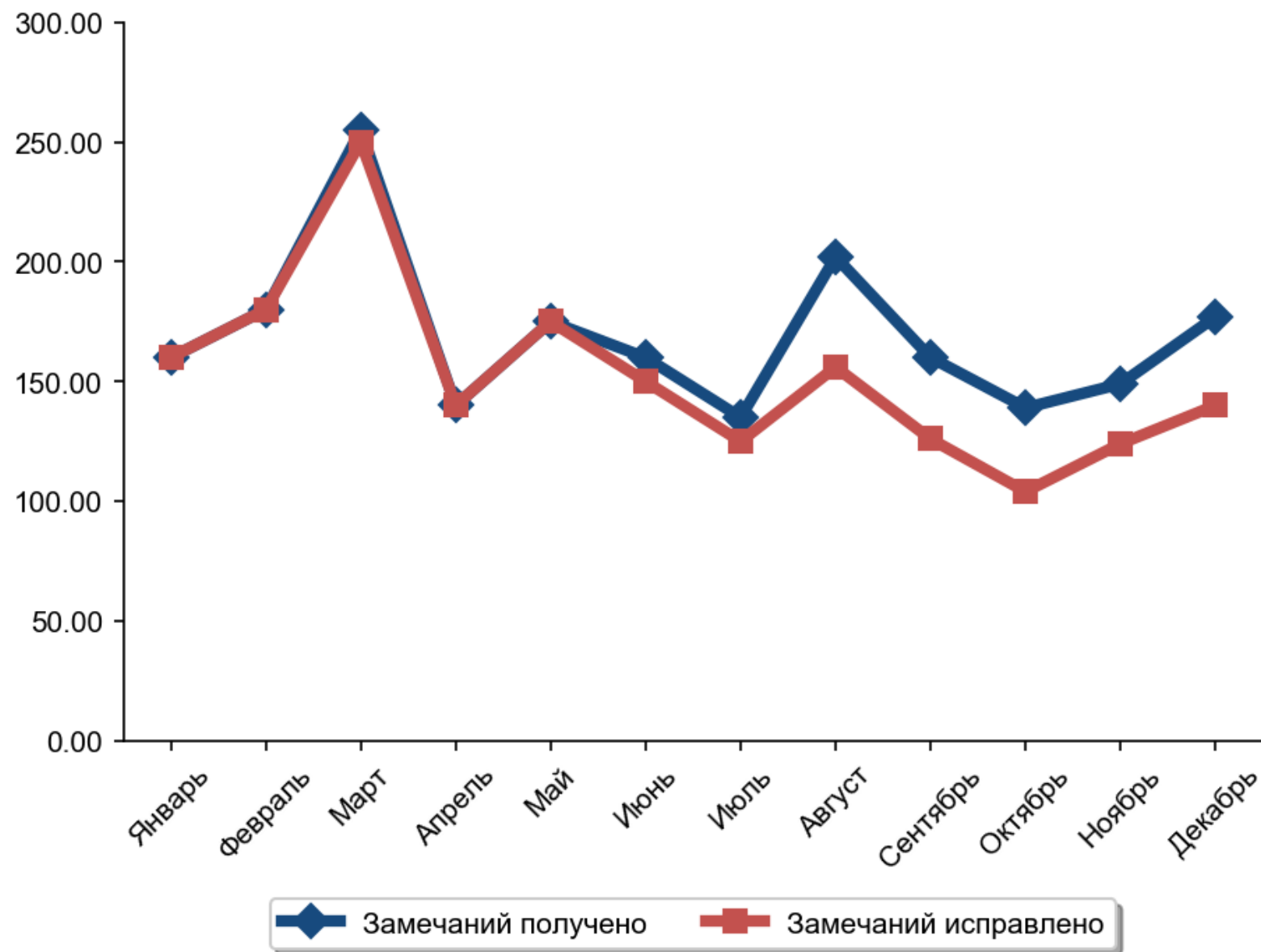
■ Segment A ■ Segment B ■ Segment C ■ Segment D

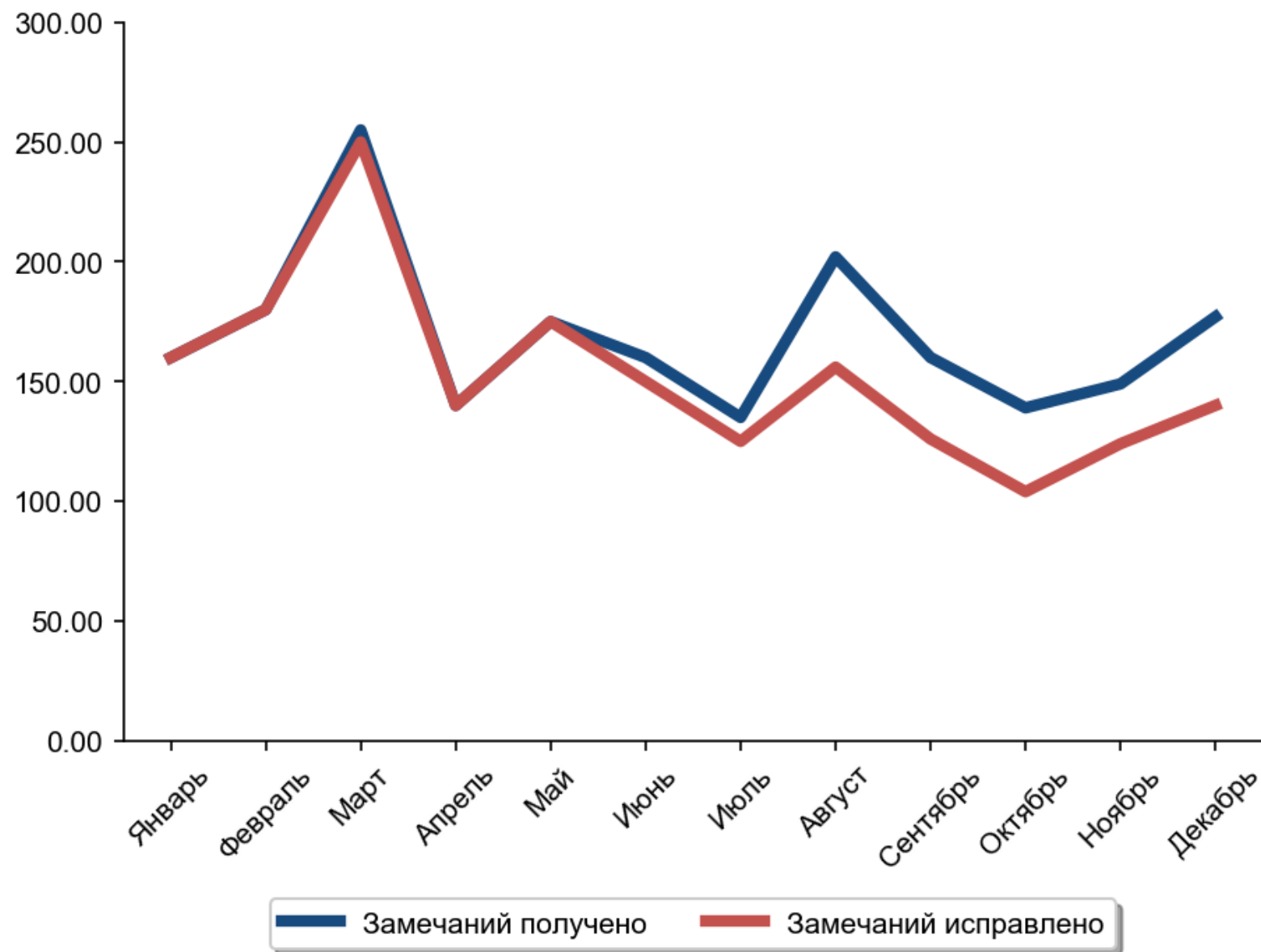


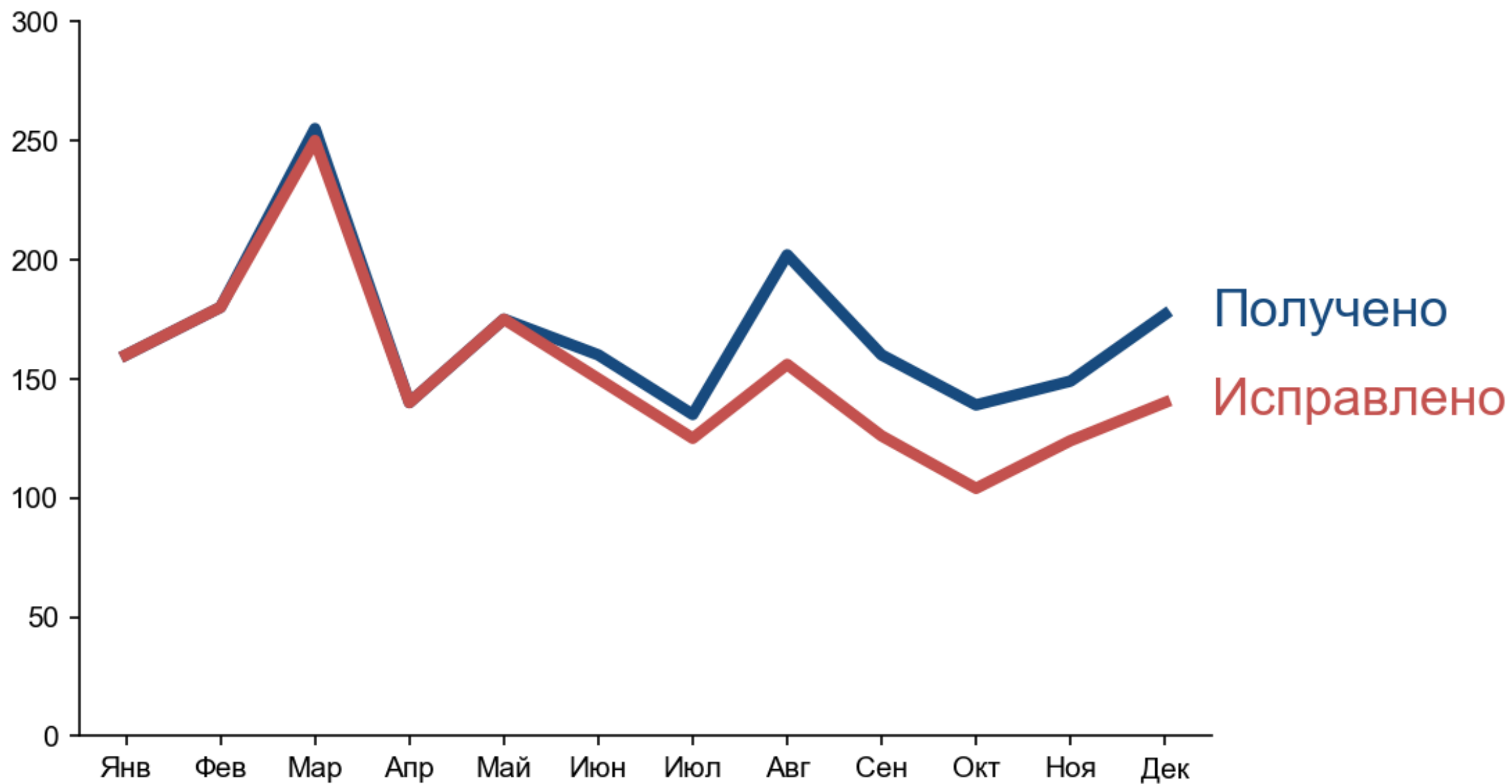
Mycop











Фокусировка

756395068473

658663037576

860372658602

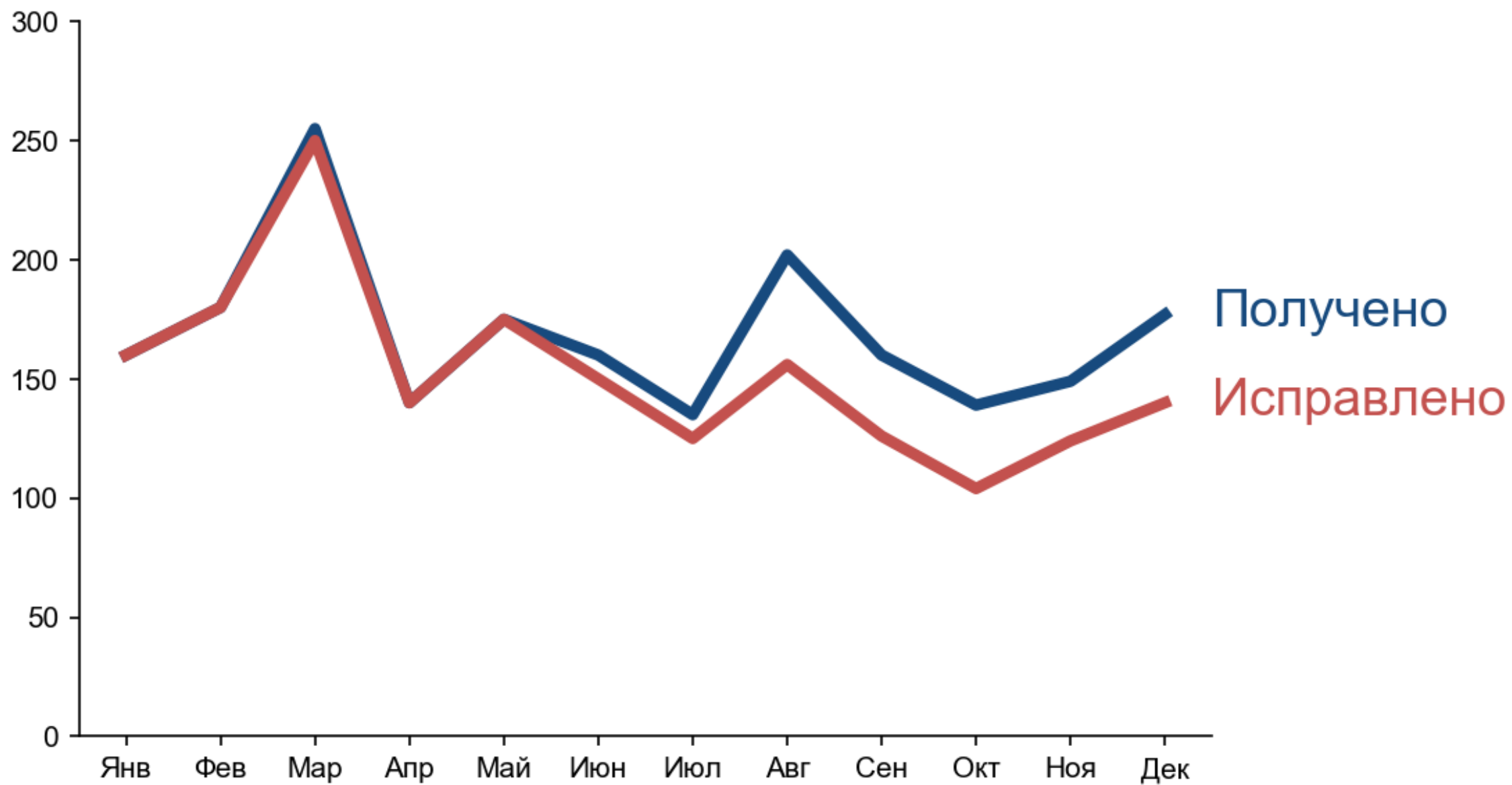
846589107830

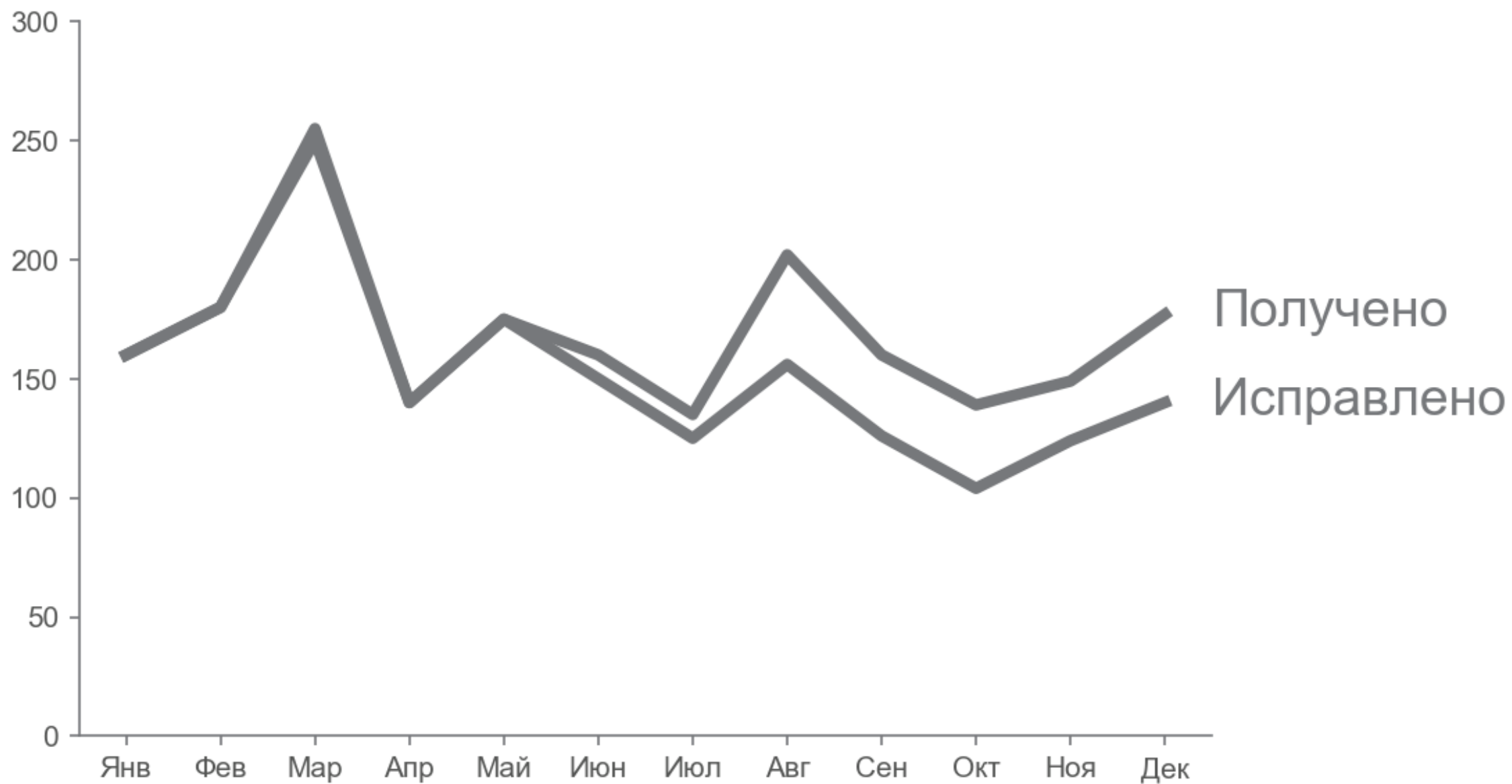
756395068473

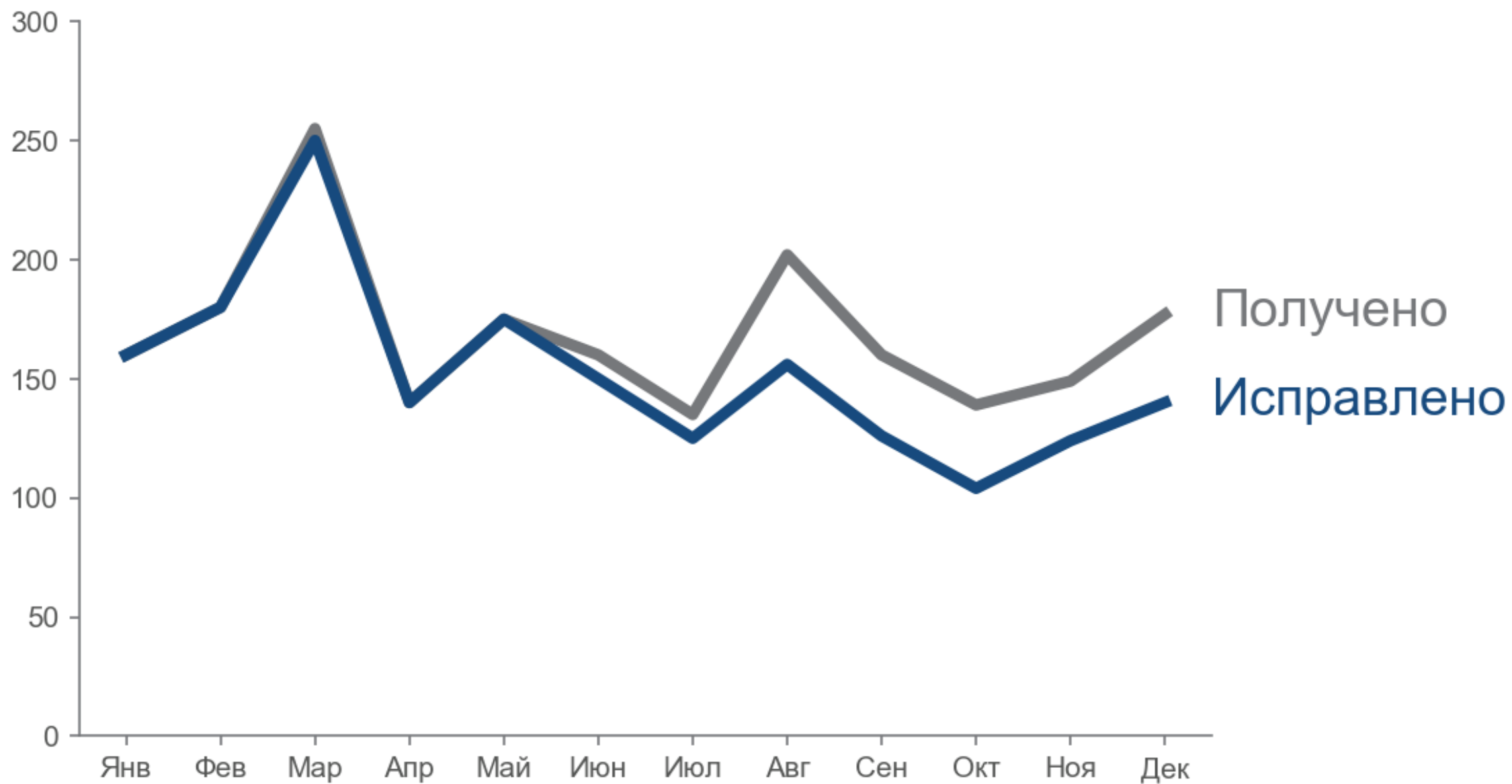
658663037576

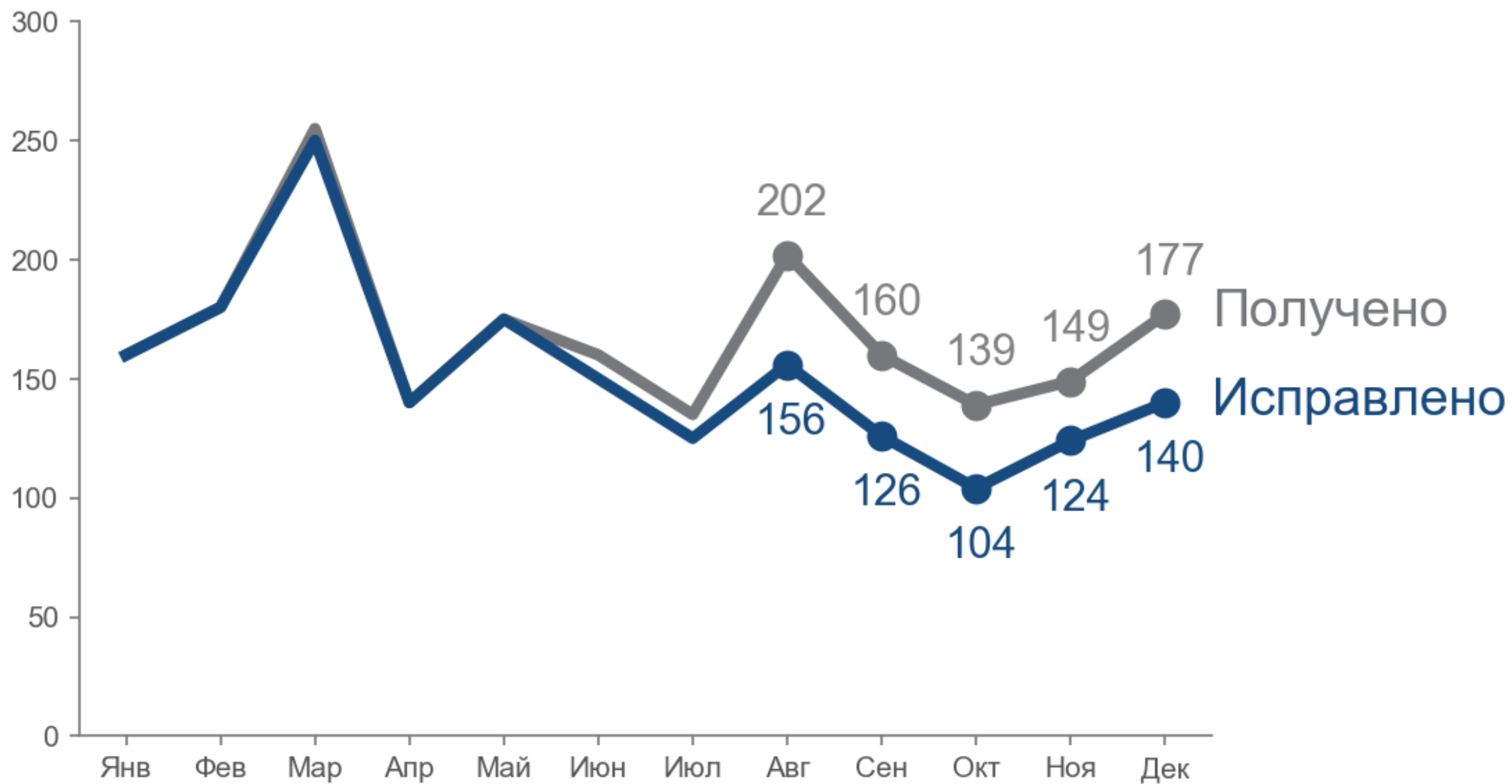
860372658602

846589107830









Дизайн

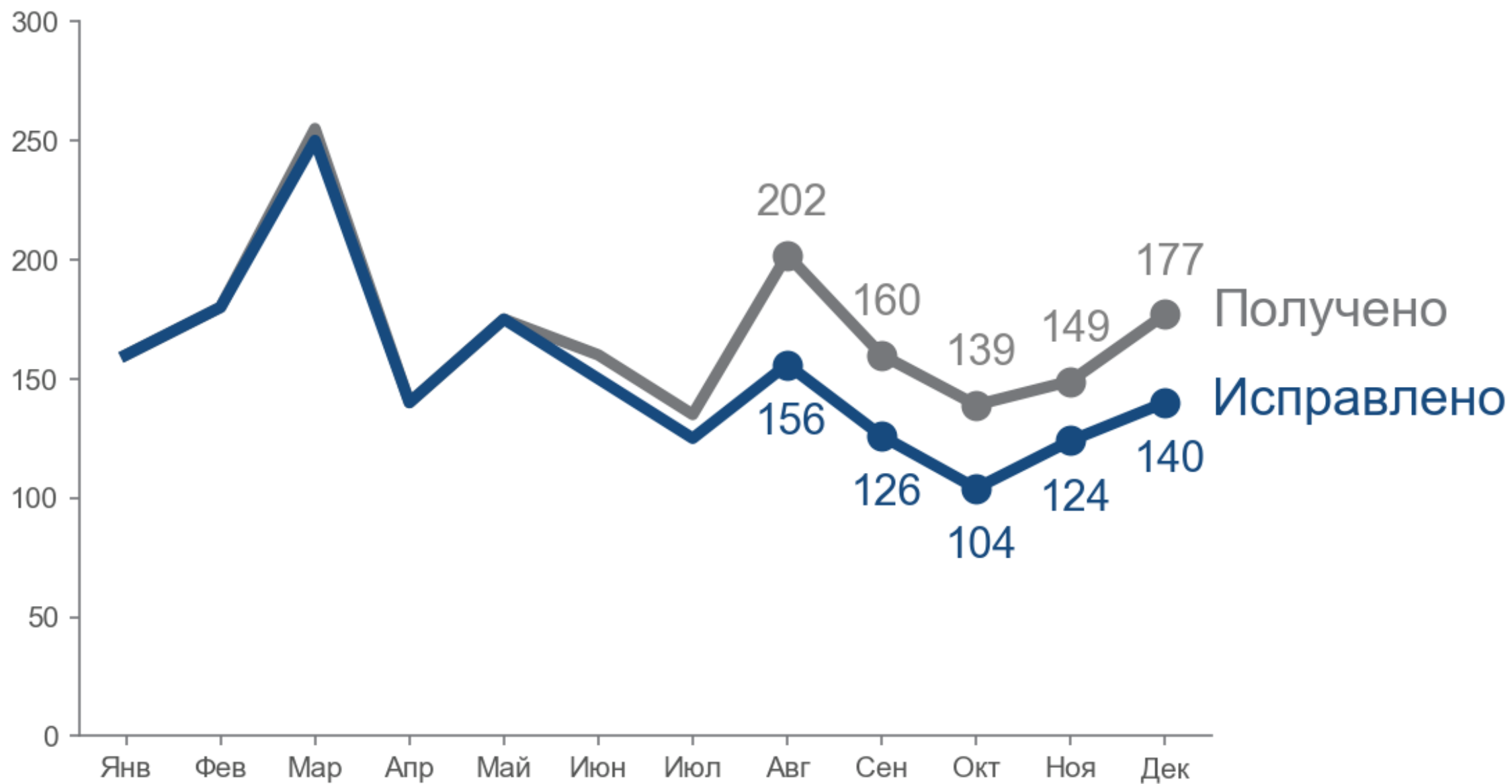
Доступность

Эстетика

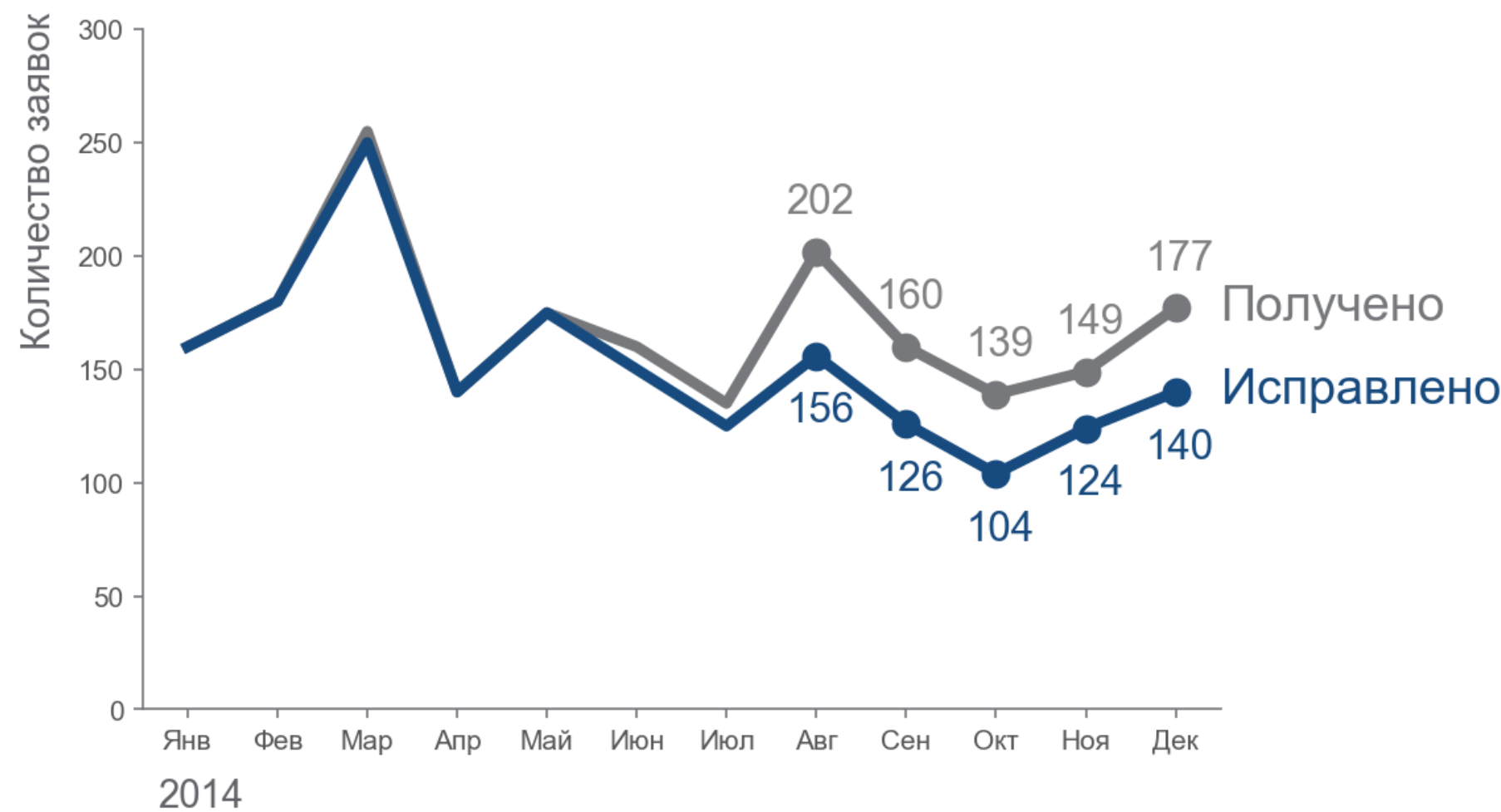
Доступность

Высококонтекстная культура

Низкоконтекстная культура



Заявки на исправления по месяцам



Данные: БД XYZ, на состояние 2015-01-01

Пожалуйста, добавьте в команду 2 разработчика

для компенсации двоих ушедших в прошлом году

Заявки на исправления по месяцам



Данные: БД XYZ, на состояние 2015-01-01 | Детальный анализ процесса обработки заявок, включающий анализ продуктивности сотрудников, может быть проведён по запросу.

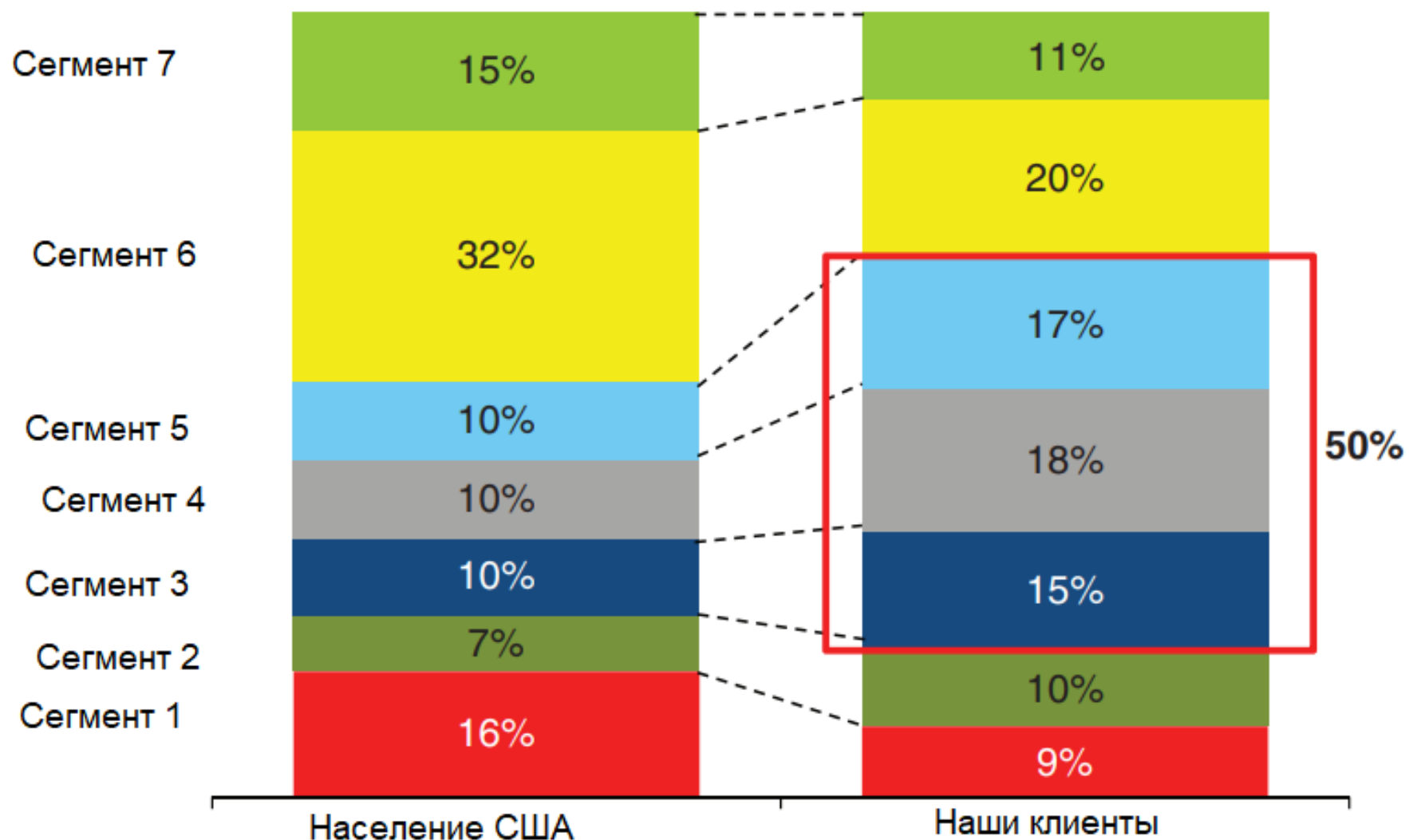
Эстетика

Выравнивайте

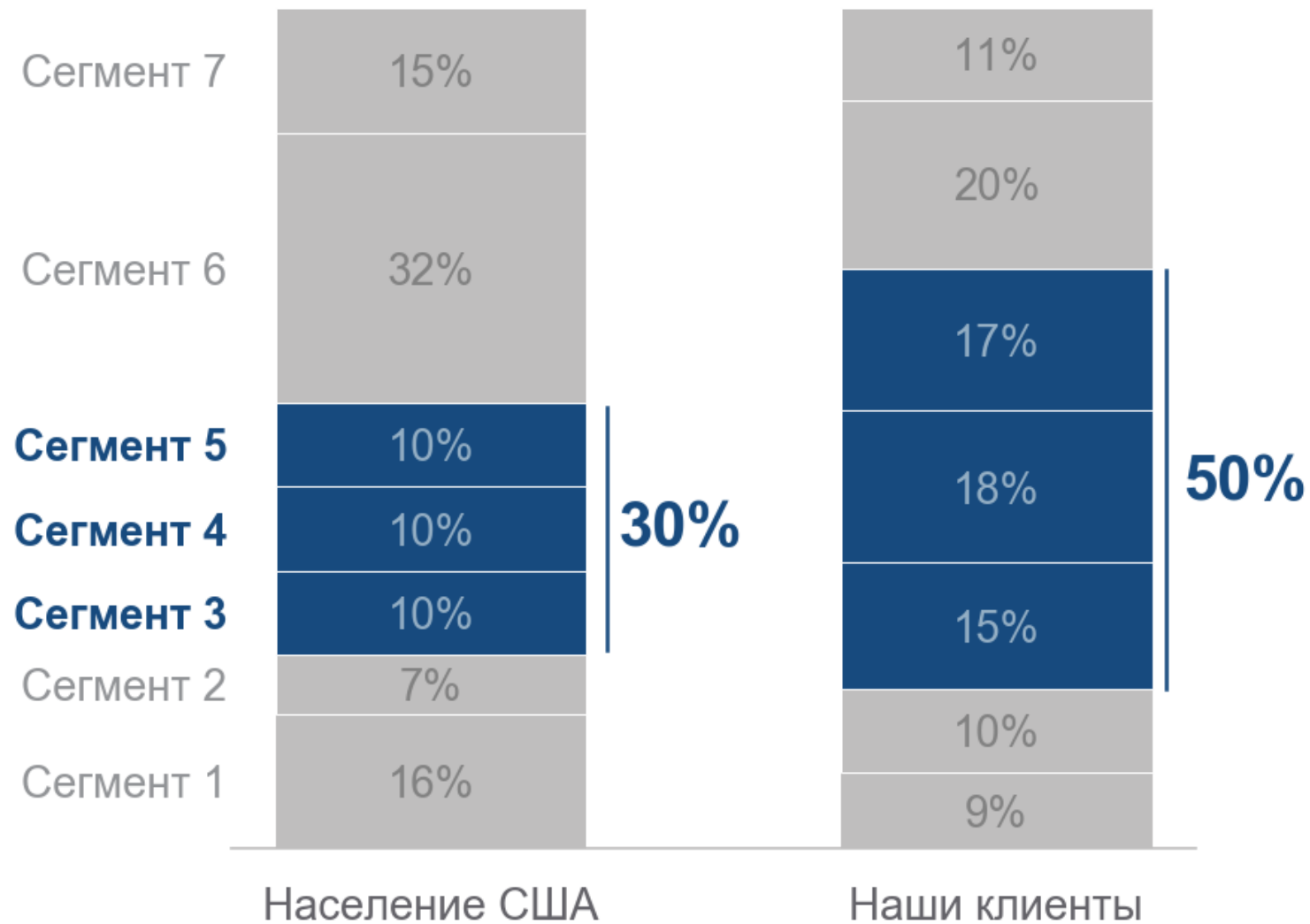
Используйте цвет разумно

Оставляйте больше пустых пространств

Распределение клиентов по сегментам



Распределение клиентов по сегментам



История

Большая Идея

Трёхминутный рассказ

Начало

«Почему меня это должно интересовать?»

Середина

«Почему я могу принять это решение?»

Конец

«Что мне нужно делать прямо сейчас?»

Стратегия рассказа

В хронологическом порядке

Начать с конца

Повторение

Поймите **контекст**

Выберите эффективную **визуализацию**

Уберите визуальный **мусор**

Сфокусируйте внимание аудитории

Проработайте **дизайн**

Составьте **историю**

Призыв к действию

Десятипальцевый метод печати

Создание эффективных визуализаций

Прочитайте книгу («Storytelling with Data»)

Выберите библиотеку (matplotlib)

Перерисуйте диаграммы

Выложите на GitHub (<https://github.com/empathy87>)