



THIS IS

# SPARTA .data-sprint

Sber Process Analytics Research Teams Association

**Команда «No more decimal»**

Андрей Супрун  
Чэмэлиинэ Александрова  
Мария Галаганова  
Александра Ковалева  
Олег Кузьмин

# Кейс: процесс обработки и доставки заказов интернет-магазина

## Описание кейса:

- магазин обрабатывает заказы с 8:00 до 23:59
- работает до последнего клиента
- входящий поток от 100 до 400 заказов ежедневно
- анализируемый период с начала октября по декабрь 2022 года
- владелец магазина ожидает, что заказ передается курьеру в течение 30 минут после сборки
- по словам владельца, в октябре и декабре были зафиксированы сбои в работе магазина, при первой итерации анализа не удалось определить характер проблем

## Цель:

Помочь владельцу интернет магазина улучшить бизнес-процесс, сократить издержки на доставку, оптимизировать работу и численность персонала

## Задачи:

- выявить классические процессные неэффективности
- рассчитать хронометраж процесса
- проверить SLA передачи заказов курьерам

# Обзор и предобработка данных

# Обзор данных

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 179030 entries, 0 to 179029
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   case            179030 non-null  int64  
1   stage           179030 non-null  object  
2   start_time      179030 non-null  datetime64[ns]
3   end_time        179030 non-null  datetime64[ns]
4   client          179030 non-null  object  
5   staff           154153 non-null  object  
6   feedback        482 non-null    object  
7   status          20248 non-null  object  
dtypes: datetime64[ns](2), int64(1), object(5)
memory usage: 10.9+ MB
```

Для исследования мы обращаемся к логам с операциями работников, в частности, обработке заказов в онлайн-магазине.

В таблице:

179 тысяч 30 строк

8 столбцов

Столбцы:

**case** - идентификаторы процессов (19988 уникальных)

**stage** - название этапа процесса (11 уникальных)

**start\_time** - временная метка начала этапа

**end\_time** - временная метка конца этапа

**client** - имя клиента

**staff** - ответственный работник

**feedback** - обратная связь

**status** - статус операции оплаты

Пропуски в значимых данных отсутствуют (status заполнен только для этапа оплаты, работник не назван при операциях, в которых он не участвует: в оформлении и отмене заказа)

Особый интерес вызывает редкая обратная связь.

# Предобработка данных

В ходе предобработки данных мы обработали логи с помощью библиотеки для process mining PM4PY, также мы добавили необходимые для дальнейшего исследования столбцы.

```
0 case
1 stage
2 start_time
3 end_time
4 client
5 staff
6 feedback
7 status
8 case:concept:name
9 concept:name
10 time:timestamp
11 @@index
12 @@case_index
13 case_total_time
14 case_stage_count
15 time_to_courrier
16 stage_time
```

Столбцы, созданные библиотекой PM4PY:

- **case:concept:name** - ID процесса
- **concept:name** - название этапа
- **time:timestamp** - временные метки конца этапа
- **@@index** - общая индексация
- **@@case\_index** - индексация процессов

Дополнительные столбцы:

- **case\_total\_time** - общая длительность процесса
  - **case\_stage\_count** - количество этапов в процессе
  - **time\_to\_courrier** - время от сборки до передачи курьеру
  - **stage\_time** - длительность этапа
- созданные

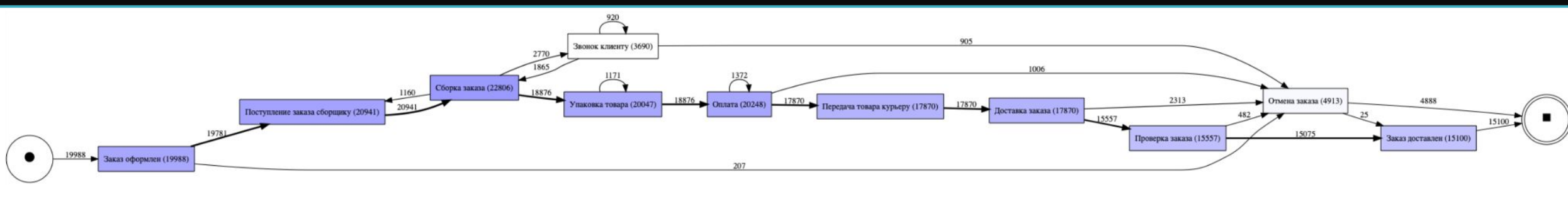
Мы преобразовали типы данных: во-первых, мы привели временные метки к типу `datetime`; во-вторых, для удобства дальнейшей работы мы привели к минутам данные типа `timedelta`.

Дополнительно мы создали две производные таблицы **success\_data** и **fail\_data**, содержащие только успешные и неуспешные процессы соответственно.



# Исследовательский анализ данных

# Модель процесса



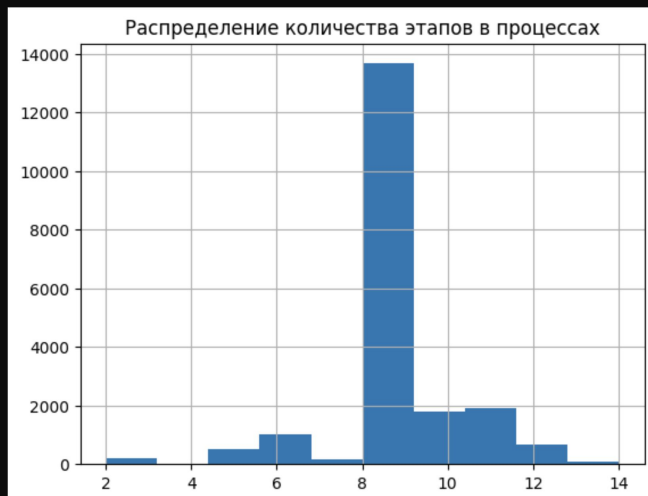
Из инструментов, предлагаемых библиотекой PM4PY, наиболее удачно отображает модель процессов заказа и доставки товаров DFG-граф.

Этот граф позволяет увидеть общую направленность всех процессов, сразу обнаружить ряд неэффективностей (зацикленности) и выявить “идеальный путь” процесса через девять неповторяющихся этапов.

## “Идеальный путь”:

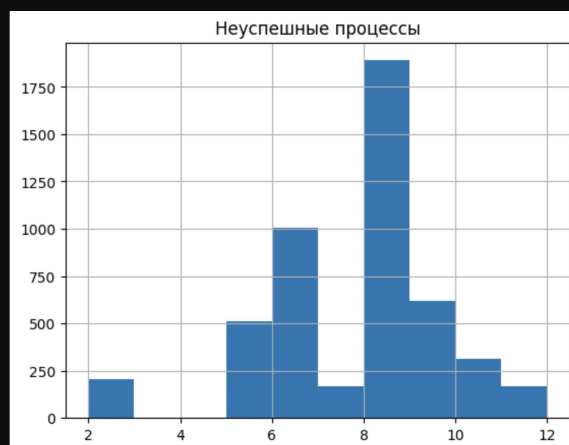
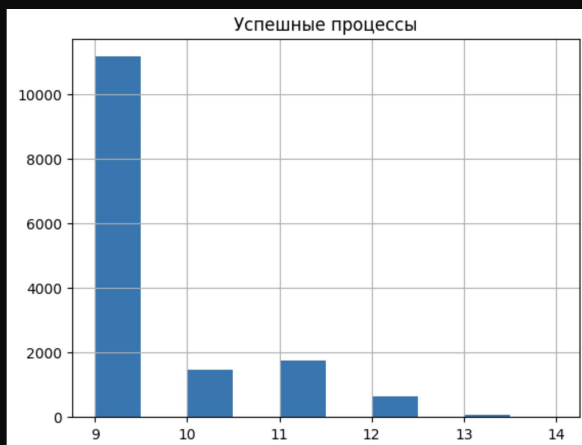
1. Заказ оформлен
2. Поступление заказа сборщику
3. Сборка заказа
4. Упаковка товара
5. Оплата
6. Передача товара курьеру
7. Доставка заказа
8. Проверка заказа
9. Заказ доставлен

# Количество этапов процесса



Количество этапов в отдельном процессе варьируется от 2 до 14. В неуспешных процессах количество изменяется от 2 до 12, в успешных процессах - от 9 до 14 этапов.

Подавляющее большинство процессов проходит за 9 этапов. Большое количество этапов вызвано повторами, т.е. нарушениями модели процесса.

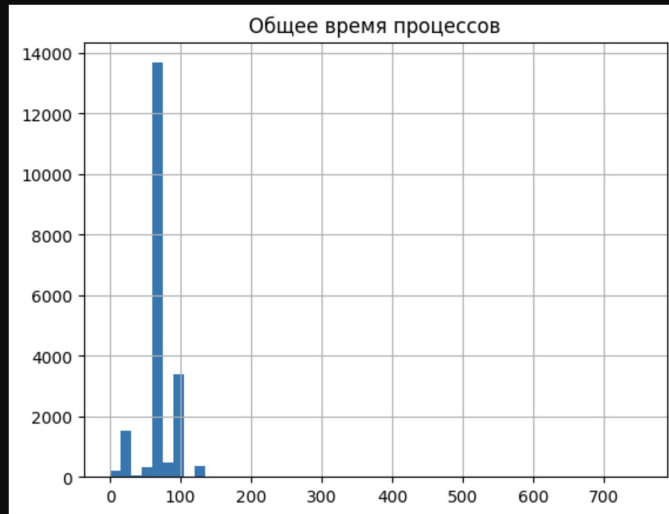


## Вывод:

Данные, полученные при моделировании процесса на DFG-графе, получили подтверждение: минимально необходимое количество этапов для успешного завершения процесса - 9 неповторяющихся этапов.



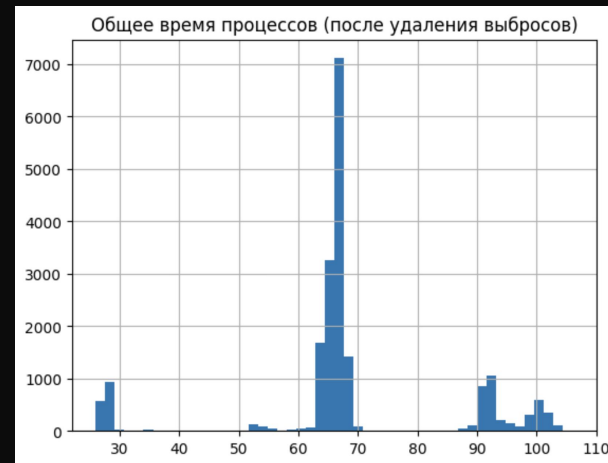
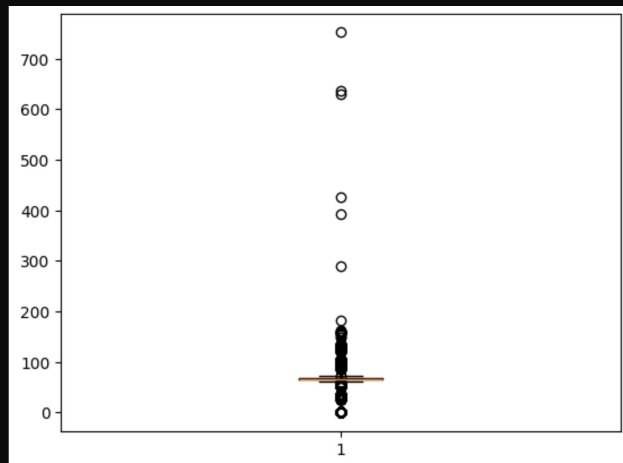
# Распределение экземпляров по длительности: удаление выбросов



При изучении длительности времени экземпляров возникает необходимость удалить выбросы (случаи длительностью от 120 до 700 минут и случаи короче 20 минут).

Удаление выбросов (575 процессов из 19988) приводит к сокращению датасета на 3%, что является допустимым.

График после удаления выбросов позволяет увидеть более точную картину распределения процессов по времени.



**Вывод:** удаление выбросов (3% данных) позволяет заметить распределение процессов по трем группам.

# Распределение экземпляров по длительности: три группы



Экземпляры по длительности можно разделить на три группы:

1. меньше 40 минут (в основном - 20-30 минут) - 1567 экземпляров (8%)
2. 40-80 минут (в основном - 60-70 минут) - 13975 экземпляров (72%)
3. больше 80 минут (в основном - 90-104 минуты) - 3871 экземпляр (20%)

Все успешные процессы входят во вторую и третью группу с преобладанием второй. Неуспешные процессы распределены между всеми тремя группами, при этом количество неуспешных процессов в первой и второй группах почти эквивалентно.



**Вывод:** большая часть процессов длится от 60 до 70 минут - это время можно принимать за усредненный стандарт, остальное считая за отклонение. Отклонение в меньшую сторону говорит о неуспешности процесса, отклонение в большую сторону - о неэффективности.

# Среднее распределение времени на отдельные этапы



Ранжируя этапы по среднему времени исполнения, можно увидеть, что большинство этапов не занимает много времени, за исключением двух - "Доставка заказа" и "Сборка заказа". При этом сумма средних значений длительности всех этапов составляет 71 минуту, что близко к средней длительности выполнения процесса (72.6 минуты).

**Вывод:** наиболее затратными по времени этапами являются "Сборка заказа" и "Доставка заказа". Именно на них необходимо обратить внимание в дальнейшем при изучении неэффективностей и поиске путей по сокращению общего времени выполнения процессов.

# SLA: работа до последнего клиента и передача заказов курьерам



## Требование работы до последнего клиента

Максимальная длительность успешного процесса - 106 минут, отмена заказа происходит максимум в срок до 104 минут после оформления. В случае отсутствия операции отмены заказ доставляется клиенту в течение менее чем 2 часов.

## Требование передачи заказов курьерам

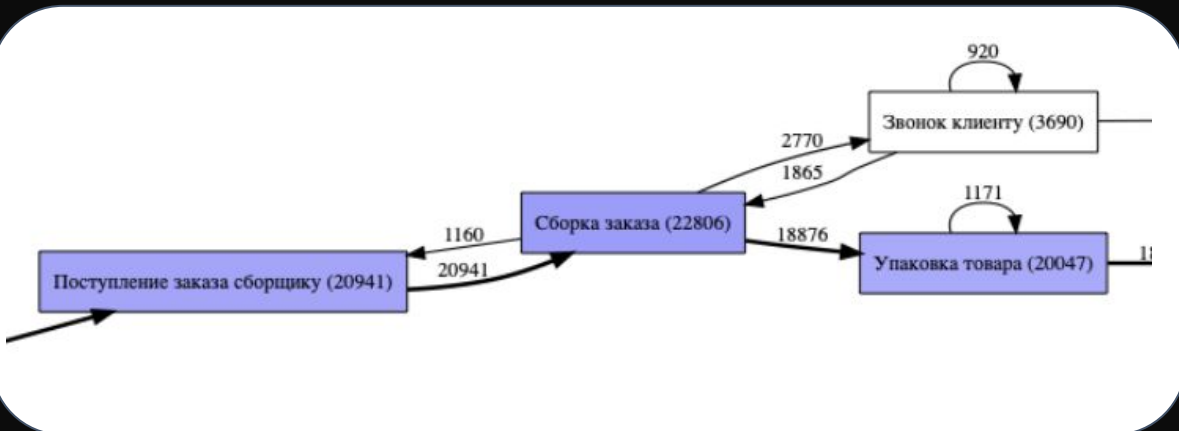
В большинстве случаев передача заказа курьеру (рассчитывается от момента завершения этапа сборки до завершения этапа передачи заказа курьеру) занимает до 10 минут, однако есть случаи, когда процесс занимает 40 и более минут. Таких случаев насчитывается 6%.

**Вывод:** требование работы до последнего клиента соблюдается для не отмененных на ранних этапах заказов, тогда как требование передачи заказов курьерам соблюдается в 94% случаев.

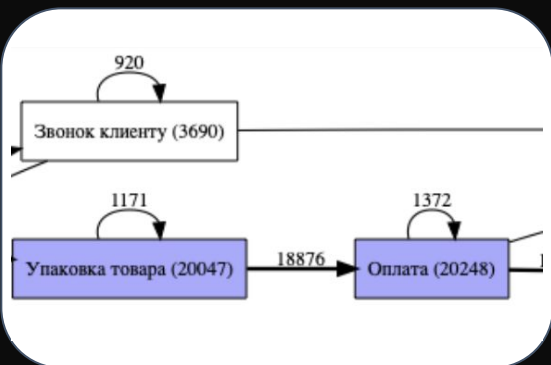
# Процессные неэффективности



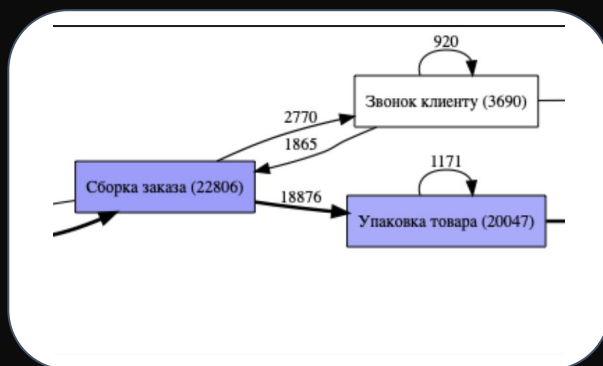
# Неэффективности: зацикленности



“Пинг-понг”, “Возврат”, “В себя”



“В себя”



“Возврат” и “В себя”

В представленных данных выделяются три вида неэффективностей группы “зацикленности”:

1. зацикленность “Возврат”
2. зацикленность “Пинг-понг”
3. зацикленность “В себя”

Они видны на представленных изображениях DFG-графа процесса.

“Возврат” возникает при возврате от этапа “Звонок клиенту” к этапу “Сбор заказа”.

“Пинг-понг” - при переходе между этапами “Сбор заказа” и “Поступление заказа сборщику”.

“В себя” - на этапах “Оплата”, “Звонок клиенту” и “Упаковка товара”.

# Неэффективности: “Возврат” и “Пинг-понг”

## “Возврат”

Зацикленность типа “возврат” возникает после этапа “Звонок клиенту”: происходит возврат на этап “Сборка заказа”, после чего процесс идет по стандартному пути

**Общее время: 40 236 минут**

**Причины возникновения:**

- \* если товар отсутствует или данные для доставки указаны неверно, необходимо связаться с клиентом, чтобы либо предложить замену товара, либо уточнить данные
- \* возврат осуществляется после подтверждения клиентом измененного заказа

**Предложения по устранению:**

- \* совершенствование автоматизированной системы, обновляющей информацию о доступных товарах на сайте
- \* автоматизация проверки наличия товара до поступления сборщику

**Оценка потенциального фин. эффекта:**

Рассчитывается как сумма длительностей первого появления операции в экземпляре, умноженная на стоимость минуты работы сотрудника.  
При стоимости 10 руб. за минуту процесса потенциальный финансовый эффект равняется 402 361 р. за три месяца работы или примерно 134 000 р. в месяц

## “Пинг-понг”

Ситуация зацикленности “пинг-понг” означает, что две операции, следующие друг за другом, повторяются, при этом зациклены и эти операции, и переход между ними. В нашем кейсе “пинг-понг” возникает между этапами “Сбор заказа” и “Поступление заказа сборщику”

**Общее время: 25 107 минут**

**Причины возникновения:**

- \* алгоритм сборки заказа не оптимизирован, вследствие чего возникает необходимость отложить операцию сборки на определенный срок

**Предложения по устранению:**

- \* оптимизация алгоритма сборки заказа

**Оценка потенциального фин. эффекта:**

Рассчитывается как сумма длительностей дублирующихся операций (за исключением последнего появления), умноженная на стоимость минуты работы сотрудника.  
При стоимости 10 руб. за минуту процесса потенциальный финансовый эффект равняется 251 070 р. за три месяца работы или примерно 83 500 р. в месяц

# Неэффективность: “В себя”

Зацикленностью “В себя” называется повторение операции два или более раз подряд. Этот тип зацикленности возникает при осуществлении трех операций: “Упаковка товара”, “Оплата”, “Звонок клиенту”.

Потенциальный финансовый эффект рассчитывается как сумма длительностей дублирующихся операций (за исключением последнего появления) на стоимость минуты работы сотрудника.

## “Упаковка товара”

**Общее время: 2 306 минут**

**Причины возникновения:**

- \* нарушения технологии проведения процесса упаковки товара

**Предложения по устранению:**

- \*дополнительная регламентация и автоматизация процесса

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 23 060 р

## “Оплата”

**Общее время: 587 минут**

**Причины возникновения:**

- \* повторные попытки платежей при ошибках платежной системы

**Предложения по устранению:**

- \* контроль работоспособности платежной системы, выбор системы с большим коэффициентом доступности

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 5 870 р

## “Звонок клиенту”

**Общее время: 2 486 минут**

**Причины возникновения:**

- \* неквалифицированное проведение первого созвона с клиентом

**Предложения по устранению:**

- \* увеличение скорости реагирования сайта на изменения ассортимента товаров, предложение автозамены на похожий товар

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 24 865 р

# Неэффективности: длительность операции

	Этап	Стандартное отклонение	Средняя на медиану
0	Доставка заказа	0.657657	0.987341
1	Сборка заказа	0.421394	0.990748

В имеющихся данных можно выделить две операции, имеющие высокую длительность - это "Сборка заказа" и "Доставка заказа". Именно они относятся к категории "нестандартный / ручной этап":

1. Операция классифицируется как длительная
2.  $0,9 < \text{Средняя длительность} / \text{медиана длительности} < 1,1$
3. Стандартное отклонение длительности операции  $> 0,3$

**Общее время: доставка заказа - 3 558 минут, сборка заказа - 1 143 минуты**

## Предложения по устранению:

- \* оптимизация маршрута курьеров для сокращения времени на доставку
- \* автоматизация и более четкая регламентация процесса сборки заказа

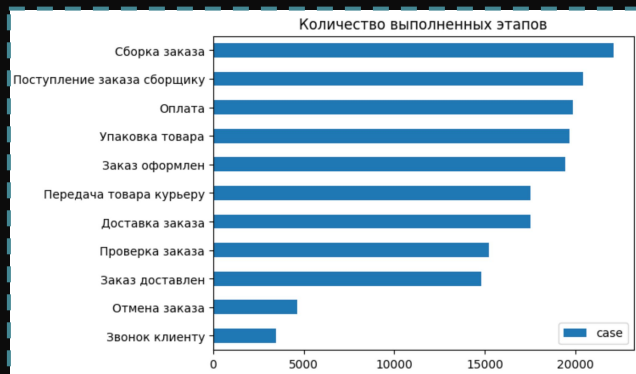
## Оценка потенциального финансового эффекта:

Потенциальный финансовый эффект рассчитывается как сумма длительностей, превышающих среднее значение для операции, умноженная на стоимость минуты работы сотрудника.

При стоимости 10 руб. за минуту процесса, потенциальный финансовый эффект за три месяца равняется:

- \* для операции "Доставка товара" - 35 585 р.
- \* для операции "Сборка заказа" - 11 434 р.

# Неэффективности: влияние на процесс



## Ошибки системы:

На подготовленном датасете выделяется 14 успешных процессов, в которых перед завершающим этапом после этапа "Доставка заказа" возникает этап "Отмена заказа". Несмотря на незначительное количество таких процессов, ситуация не может остаться без внимания, потому что свидетельствует об ошибке в автоматизированной системе, которая потенциально может привести к большому сбою.

Также ошибками системы являются операции ошибок в системе оплаты (они приводят к заикливаннию операции "Оплата" и посчитаны в разделе "Заикликованность: в себя")

## Нерегулярные операции:

Этап "Звонок клиенту" - нерегулярный ( 2512 операций, 12.9%, без учета заикликованностей - 1617 операций)

**Общее время: 4 615 минут**

**Причины возникновения:** недостаточно эффективные механизмы обработки информации о клиентах на сайте, недостаточно быстрое обновление информации о доступных и недоступных для заказа товарах

**Предложения по устранению:**

- \* Доработка и настройка системы мониторинга
- \* Уменьшение времени обновления системы
- \* Интеграция информационных систем, внедрение RPA для минимизации ручных действий

**Оценка потенциального финансового эффекта:**

Потенциальный финансовый эффект рассчитывается как сумма длительности нерегулярных операций за исключением случаев, признанных заикликованностью, умноженная на стоимость минуты работы сотрудника.

При стоимости 10 руб. за минуту процесса потенциальный финансовый эффект за три месяца равняется 46 155 р



# Неэффективности: неуспех



## Варианты неуспешного завершения экземпляров:

1. После звонка клиенту:  
905 случаев - 19.5%
2. После проверки:  
447 случаев - 9.6%
3. При оплате:  
1006 случаев - 21.7%
4. Во время доставки:  
2267 - 49%

## Причины возникновения ситуации неуспеха :

1. Отсутствие заказанного товара на складе и невозможность предоставить удовлетворяющую клиента замену
2. Некачественная сборка заказа (поврежденный товар, ошибочно отправленный товар)
3. Ошибка обработки платежа платежной системой
4. Требуется наибольшего внимания, т.к. отмена происходит в процессе доставки, что свидетельствует о нарушении процесса

## Оценка потенциального финансового эффекта:

Потенциальный финансовый эффект рассчитывается как сумма длительностей экземпляров без операции успеха на стоимость минуты работы сотрудника

# Неэффективности: успех

## “Звонок”

**Общее время неуспешных экземпляров: 30 041 минута**

**Предложения по устранению:**

\* совершенствование системы обновления информации о доступности товаров на сайте

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 300 410 р

## “Проверка”

**Общее время неуспешных экземпляров: 32 525 минут**

**Предложения по устранению:**

\* совершенствование работы по контролю качества сборки товаров

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 325 254 р

## “Оплата”

**Общее время неуспешных экземпляров:  
32 865 минут**

**Предложения по устранению:**

\* проверка корректности работы системы оплаты

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 328 655 р

## “Доставка”

**Общее время неуспешных экземпляров:  
160 227 минут**

**Предложения по устранению:**

\* требует наибольшего внимания, т.к. отмена происходит в процессе доставки, что свидетельствует о нарушении процесса

**Оценка потенциального фин. эффекта:**

Потенциальный финансовый эффект за три месяца равняется 1 602 276 р

# Неэффективности: вывод

В ходе исследования удалось выявить следующие процессные неэффективности:

1. три типа неэффективностей группы "Зацикленностей":
  - возврат на этап "Сборки заказа" после "Звонка клиенту"
  - "пинг-понг" между этапами "Сборка заказа" и "Поступление заказа сборщику"
  - "в себя" на этапах "Звонок клиенту", "Оплата" и "Упаковка товара"
2. неэффективность группы "Длительность операции":
  - этапы "Доставка заказа" и "Сборка заказа" классифицируются как "нестандартные / ручные этапы"
3. два типа неэффективностей группы "Влияние на процесс":
  - ошибки системы: одновременность наличия этапа "Отмена заказа" и этапа "Заказ доставлен", ошибки на этапе "Оплата"
  - нерегулярные операции: этап "Звонок клиентам"

## Пути формирования неуспешных процессов:

1. отмена заказа после звонка клиенту
2. отмена заказа после проверки
3. отмена заказа на этапе оплаты
4. отмена заказа во время доставки

## Основные рекомендации:

- проверка качества, надежности и отказоустойчивости используемой платежной системы
- регламентация работы, совершенствование системы обновления количества товаров в наличии на сайте магазина
- регламентация и автоматизация этапа
- усиление системы контроля качества

# Итоги: этапы исследования

В рамках исследования мы провели работу по изучению процессов обработки и доставки заказов интернет-магазина.

## На этапе обзора и предобработки данных:

- получили логи процессов обработки и доставки заказов (таблица из 8 столбцов и 179 тыс 30 строк)
- обработали логи с помощью библиотеки PM4PY, добавили новые столбцы (общая длительность процессов, количество этапов, время от завершения сборки до передачи товара курьеру, длительность каждого этапа)
- преобразовали типы данных в корректные форматы (datetime, минуты вместо timedelta)
- создали две дополнительные таблицы (только успешные и только неуспешные процессы)

## На этапе исследовательского анализа:

- построили общую модель прохождения процессов в виде DFG
- выделили выбросы в данных и устранили их (3% от датасета)
- проследили хронометраж, разделив процессы на три группы - 25-30, 60-70 и 90-100 минут
- выделили "идеальный" срок протекания процесса - 9 неповторяющихся этапов за 60-70 минут
- проверили выполнение SLA: требование по "работе до последнего клиента" соблюдается, а требование по передаче заказа курьеру в течение 40 минут соблюдается в 94% случаев

## Выделение процессных неэффективностей:

- выделяется большая группа "заикленностей": возврат на этап сборки заказа, "пинг-понг" между сборкой и поступлением заказа сборщику, заикленности "в себя" на этапах "Звонок клиенту", "Оплата" и "Упаковка товара"
- две операции, а именно "Сборка заказа" и "Доставка заказа", классифицируются как длительные типа "нестандартные / ручные операции"
- выделяются также такие неэффективности, как "ошибки системы" (этап "Отмена заказа" при наличии этапа "Заказ доставлен") и "нерегулярные операции" ("Звонок клиенту")

# Итоги: потенциальный эффект и рекомендации

## Оценка потенциального эффекта:

- суммарный финансовый эффект от неуспешных экземпляров потенциально составляет 2 556 596 р. за три месяца
- больше половины указанной суммы приходится на отмену заказов во время этапа доставки
- оставшаяся часть равномерно распределяется между этапами проверки, оплаты и звонков клиентам
- из остальных неэффективностей наиболее крупными и потенциально опасными оказываются неэффективности заикливания: "Возврат" и "Пинг-понг", сосредоточенные вокруг этапа "Сборка заказа" - суммарно потенциальный финансовый эффект составляет 653 431 р.
- эффект от остальных неэффективностей за три месяца работы на порядок меньше (десятки тысяч)

## Основные рекомендации:

- для этапа "Оплата" - проверка качества, надежности и отказоустойчивости используемой платежной системы
- для этапа "Звонок клиенту" - регламентация работы, совершенствование системы обновления количества товаров в наличии на сайте магазина
- для этапов "Сборка заказа" и "Доставка заказа" - регламентация и автоматизация этапа
- для этапа "Упаковка товара" - усиление системы контроля качества
- в целом, в первую очередь необходимо сосредоточить внимание на совершенствовании трех этапов работы:
  1. этап "Сборка заказа" - требует регламентации деятельности и обучения работников, чтобы сократить количество заикливаний и повторений процедур, сокращение длительности операции (стремиться к 10 минутам)
  2. этап "Доставка заказа" требует большего контроля за деятельностью курьеров, своевременностью доставки, т.к. является потенциально самым финансово опасным
  3. этап "Звонок клиенту" - это этап, который необходимо минимизировать за счет совершенствования автоматизации процессов сбора информации о клиенте, о заказе и о наличии товаров



# Контакты

Андрей Супрун  
Чэмэлиинэ Александрова  
Мария Галаганова  
Александра Ковалева  
Олег Кузьмин

souprunandrew@gmail.com  
aleksandrova.chemeliine@mail.ru  
ms.galaganova@gmail.com  
akkovaleva14@gmail.com  
kuzmin.oleg.rus@gmail.com