Анализ эффективности работы компании Pens and Pencils и рекомендации по масштабированию бизнеса

Выполнила: Горлова Кристина Олеговна

# ЗАДАЧА

Проанализировать работу компании с точки зрения ее эффективности и дать рекомендации по масштабированию бизнеса.

# ФОРМАЛИЗОВАННАЯ ЗАДАЧА

1. Оценить динамику продаж и распределение выручки по товарам.
2. Составить портрет клиента, выяснить, какие клиенты приносят больше всего выручки.
3. Проконтролировать логистику компании: определить, все ли заказы доставляются в срок и в каком штате лучше открыть офлайн-магазин.

# Определение эффективности продаж

## 1.1. Определение динамики продаж

Для расчета выручки, полученной компанией по месяцам был составлен следующий запрос:

SELECT

DATE\_TRUNC ('MONTH', sd.order\_date)::date date,

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue

FROM

sql.store\_carts sc

JOIN sql.store\_products sp

ON sc.product\_id=sp.product\_id

JOIN sql.store\_delivery sd

ON sc.order\_id=sd.order\_id /\*через order\_id\*/

GROUP BY 1

ORDER BY 1

Детализация и комментарии:

SELECT

DATE\_TRUNC ('MONTH', sd.order\_date)::date date, /\*столбец, выводящий округление даты заказа до месяца при помощи функции DATE\_TRUNC и типизация полученных данных в дату\*/

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue /\* столбец, выводящий сумму выручки по заказам за вычетом скидки, округленной до целого числа при помощи функции ROUND\*/

FROM

sql.store\_carts sc /\*из таблицы со списком товаров в заказах\*/

JOIN sql.store\_products sp /\*соединение со списком продуктов\*/

ON sc.product\_id=sp.product\_id /\*через product\_id\*/

JOIN sql.store\_delivery sd /\*соединение со списком доставок\*/

ON sc.order\_id=sd.order\_id /\*через order\_id\*/

GROUP BY 1/\*группировка данных по дате заказа\*/

ORDER BY 1/\*сортировка полученных данных по дате заказа\*/

По результатам, полученным из запроса была построен график, отражающий динамику выручки по месяцам.



Можно отметить итоговый рост выручки в сравнении с первым месяцем. Также, наблюдается сезонность: в январе, феврале и октябре объем выручки снижается, а в марте, сентябре и ноябре объем выручки увеличивается. Это может быть связано с началом офисного сезона в сентябре и черной пятницей в ноябре. Сложно объяснить рост продаж в марте, возможно это связано с особенностями американского рынка. Для преодоления сезонного спада продаж можно увеличить рекламную активность и установить дополнительные скидки, особенно для постоянных клиентов.

## 1.2. Определение наиболее выгодных категории и подкатегории продаж,которые приносят больше всего выручки

Для определения категорий и подкатегорий товаров, приносящих наибольший объем выручки был составлен следующий запрос:

SELECT

sp.category,

sp.subcategory,

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue

FROM

sql.store\_carts sc

JOIN sql.store\_products sp

ON sc.product\_id=sp.product\_id

JOIN sql.store\_delivery sd

ON sc.order\_id=sd.order\_id

GROUP BY 1,2

ORDER BY 3 DESC

Детализация и комментарии:

SELECT

sp.category, /\*столбец, выводящий категорию товаров\*/

sp.subcategory, /\*столбец, выводящий подкатегорию товаров\*/

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue /\* столбец, выводящий сумму выручки по категориям и подкатегориям за вычетом скидки через функцию SUM, округленной до целого числа при помощи функции ROUND\*/

FROM

sql.store\_carts sc /\*из таблицы со списком товаров в заказах\*/

JOIN sql.store\_products sp /\*соединение со списком продуктов\*/

ON sc.product\_id=sp.product\_id /\*через product\_id\*/

JOIN sql.store\_delivery sd /\*соединение со списком доставок\*/

ON sc.order\_id=sd.order\_id /\*через order\_id\*/

GROUP BY 1,2 /\*группировка данных по категории и подкатегории товаров\*/

ORDER BY 3 DESC /\*сортировка полученных данных по выручке в обратном порядке\*/

Больше всего выручки принесли следующие товары:

* Furniture Chairs - 235,318 $;
* Technology Phones - 221,110 $;
* Office Supplies Storage - 179,736 $.

Меньше всего выручки принесли следующие товары:

* Office Supplies Fasteners - 2,459 $;
* Office Supplies Labels - 10,345 $;
* Office Supplies Envelopes - 13,809 $.

На основании полученных данных построена круговая диаграмма:



Из данной диаграммы можно сделать вывод что, офисные кресла, телефоны и органайзеры приносят 16%, 15% и 12% выручки соответственно.

## 1.3. Определение наиболее продаваемых товаров

Для определения 25 наиболее продаваемых товаров был составлен следующий запрос:

WITH product\_revenue AS

(SELECT

sp.product\_nm,

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount)),2) revenue,

SUM(sc.quantity) quantity

FROM

sql.store\_carts sc

JOIN sql.store\_products sp

ON sc.product\_id=sp.product\_id

JOIN sql.store\_delivery sd

ON sc.order\_id=sd.order\_id

GROUP BY 1

ORDER BY 1)

SELECT

product\_nm,

revenue,

quantity,

ROUND ((revenue / SUM(revenue) over () \* 100),2) percent\_from\_total FROM

product\_revenue

ORDER BY 2 DESC

LIMIT 25

Детализация и комментарии:

WITH product\_revenue AS /\*ключевое слово WITH и наименование подзапроса\*/

(SELECT /\*начало подзапроса\*/

sp.product\_nm, /\*столбец, выводящий наименование товара\*/

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount)),2) revenue, /\*столбец, выводящий выручку по товарам за вычетом скидки через функцию SUM, округленной до двух знаков после запятой при помощи функции ROUND\*/

SUM(sc.quantity) quantity /\*столбец, выводящий сумму товаров по категориям через функцию SUM\*/

FROM

sql.store\_carts sc /\*из таблицы со списком товаров в заказах\*/

JOIN sql.store\_products sp /\*соединяем с таблицей со списком продуктов\*/

ON sc.product\_id=sp.product\_id /\*через product\_id\*/

JOIN sql.store\_delivery sd /\*соединяем со списком доставок\*/

ON sc.order\_id=sd.order\_id /\*через order\_id\*/

GROUP BY 1 /\*группировка данных по названию продукта\*/

ORDER BY 1) /\*сортировка данных по названию продукта и окончание подзапроса\*/

SELECT

product\_nm, /\*столбец, выводящий наименование товара\*/

revenue, /\*столбец, выводящий выручку\*/

quantity, /\*столбец, выводящий количество товаров\*/

ROUND ((revenue / SUM(revenue) over () \* 100),2) percent\_from\_total /\*столбец, выводящий долю продукта в общей выручке в процентах, округленной до двух знаков после запятой при помощи функции ROUND следующим образом: выручку от продукта делим на общую выручку, рассчитанную при помощи оконной функции через функцию SUM, в которой окном является вся таблица\*/

FROM

product\_revenue /\*из подзапроса product\_revenue\*/

ORDER BY 2 DESC /\*сортировка полученных данных по выручке с продукта в обратном порядке\*/

LIMIT 25 /\*ключевое слово LIMIT, выводящее только то количество строк, которое указано далее, в данном случае 25 строк\*/

Результат данного запроса представлен в следующей таблице:

| **product\_nm** | **revenue** | **quantity** | **percent\_from\_total** |
| --- | --- | --- | --- |
| Canon imageCLASS 2200 Advanced Copier | 36 959,82 | 20 | 2,56 |
| HON 5400 Series Task Chairs for Big and Tall | 15 309,53 | 39 | 1,06 |
| 3D Systems Cube Printer, 2nd Generation, Magenta | 14 299,89 | 11 | 0,99 |
| High Speed Automatic Electric Letter Opener | 13 624,21 | 11 | 0,94 |
| Martin Yale Chadless Opener Electric Letter Opener | 13 325 | 22 | 0,92 |
| Cisco TelePresence System EX90 Videoconferencing Unit | 11 319,24 | 6 | 0,78 |
| Hewlett Packard LaserJet 3310 Copier | 11 303,69 | 38 | 0,78 |
| Samsung Galaxy Mega 6.3 | 11 154,87 | 39 | 0,77 |
| Canon PC1060 Personal Laser Copier | 9 295,83 | 19 | 0,64 |
| HP Designjet T520 Inkjet Large Format Printer - 24" Color | 9 187,5 | 12 | 0,64 |
| Global Troy Executive Leather Low-Back Tilter | 9 082,87 | 32 | 0,63 |
| Honeywell Enviracaire Portable HEPA Air Cleaner for 17' x 22' Room | 9 043,55 | 40 | 0,63 |
| Tennsco 6- and 18-Compartment Lockers | 8 740,17 | 43 | 0,6 |
| Plantronics CS510 - Over-the-Head monaural Wireless Headset System | 8 657,89 | 35 | 0,6 |
| SAFCO Arco Folding Chair | 8 100,95 | 53 | 0,56 |
| Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish | 7 805,48 | 24 | 0,54 |
| Apple iPhone 5 | 7 798 | 24 | 0,54 |
| Tennsco Double-Tier Lockers | 7 704,86 | 47 | 0,53 |
| Ativa V4110MDD Micro-Cut Shredder | 7 699,89 | 11 | 0,53 |
| Hon Deluxe Fabric Upholstered Stacking Chairs, Rounded Back | 7 446,44 | 51 | 0,51 |
| Tennsco Single-Tier Lockers | 7 326,59 | 28 | 0,51 |
| Office Star - Professional Matrix Back Chair with 2-to-1 Synchro Tilt and Mesh Fabric Seat | 7 019,5 | 28 | 0,49 |
| Zebra ZM400 Thermal Label Printer | 6 965,7 | 6 | 0,48 |
| Hot File 7-Pocket, Floor Stand | 6 510,77 | 49 | 0,45 |
| Bretford Rectangular Conference Table Tops | 6 497,82 | 46 | 0,45 |

На основании полученных результатов можно сделать вывод о том, что больше всего выручки компании приносит продажа копировального аппарата Canon imageCLASS 2200 Advanced Copier, офисного кресла HON 5400 Series Task Chairs for Big and Tall и принтера 3D Systems Cube Printer, 2nd Generation, Magenta.

# Составление портрета клиента

## 2.1. Определение соотношения категории клиентов В2В и В2С

Для определения объемов выручки, которую приносят В2В и В2С клиенты был составлен следующий запрос:

SELECT

s\_cust.category,

COUNT(DISTINCT s\_cust.cust\_id) cust\_cnt,

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue

FROM

sql.store\_carts sc

JOIN sql.store\_products sp

ON sc.product\_id=sp.product\_id

JOIN sql.store\_delivery sd

ON sc.order\_id=sd.order\_id

JOIN sql.store\_customers s\_cust

ON sd.cust\_id = s\_cust.cust\_id

GROUP BY 1

ORDER BY 3 DESC

Детализация и комментарии:

SELECT

s\_cust.category, /\*столбец, выводящий категорию товара\*/

COUNT(DISTINCT s\_cust.cust\_id) cust\_cnt, /\*столбец, выводящий количество уникальных товаров по cust\_id через функцию COUNT\*/

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount))) revenue /\*столбец, выводящий выручку по товарам за вычетом скидки через функцию SUM, округленной до целого при помощи функции ROUND\*/

FROM

sql.store\_carts sc /\*из таблицы со списком товаров в заказах\*/

JOIN sql.store\_products sp /\*соединяем с таблицей со списком продуктов\*/

ON sc.product\_id=sp.product\_id /\*через product\_id\*/

JOIN sql.store\_delivery sd /\*соединяем с таблицей со списком доставок\*/

ON sc.order\_id=sd.order\_id /\*через order\_id\*/

JOIN sql.store\_customers s\_cust /\*соединяем с таблицей со списком клиентов через таблицу со списком доставок\*/

ON sd.cust\_id = s\_cust.cust\_id /\*через cust\_id\*/

GROUP BY 1 /\*группировка данных по категории товара\*/

ORDER BY 3 DESC /\*сортировка данных по количеству товаров\*/

Результат выполнения данного запроса выглядит следующим образом:

| category | cust\_cnt | revenue |
| --- | --- | --- |
| Corporate | 645 | 1172009 |
| Consumer | 148 | 274148 |

Можно сделать вывод о том, что корпоративных клиентов в несколько раз больше, также, выручка от них в несколько раз больше.

## 2.2. Определение динамики новых В2В клиентов по месяцам

Для определения динамики появления новых В2В клиентов по месяцам был составлен следующий запрос:

WITH new\_cust AS

(SELECT

MIN (DATE\_TRUNC ('MONTH', sd.order\_date))::date "month",

COUNT (DISTINCT s\_cust.cust\_id) cust\_cnt

FROM

sql.store\_customers s\_cust

JOIN sql.store\_delivery sd

ON s\_cust.cust\_id=sd.cust\_id

GROUP BY s\_cust.cust\_id, s\_cust.category

HAVING s\_cust.category = 'Corporate'

ORDER BY 1)

SELECT

"month",

SUM(cust\_cnt) cust\_cnt

FROM new\_cust

GROUP BY 1

ORDER BY 1

Детализация и комментарии:

WITH new\_cust AS /\*ключевое слово WITH и наименование подзапроса\*/

(SELECT/\*начало подзапроса\*/

MIN (DATE\_TRUNC ('MONTH', sd.order\_date))::date "month", /\*столбец, содержащий минимальную дату, в которую клиент совершил заказ, определяемый функцией MIN, с округлением даты заказа до месяца при помощи функции DATE\_TRUNC и типизация полученных данных в date\*/

COUNT (DISTINCT s\_cust.cust\_id) cust\_cnt /\*столбец, содержащий количество уникальных cust\_id на выведенную дату через функцию COUNT \*/

FROM

sql.store\_customers s\_cust /\*из таблицы со списком клиентов\*/

JOIN sql.store\_delivery sd /\*соединяем с таблицей со списком доставок\*/

ON s\_cust.cust\_id=sd.cust\_id /\*через cust\_id\*/

GROUP BY s\_cust.cust\_id, s\_cust.category /\*группировка данных по cust\_id и категории товара\*/

HAVING s\_cust.category = 'Corporate' /\*фильтрация данных по категории клиента равное ‘Corporate’\*/

ORDER BY 1) /\*сортировка данных по минимальной дате заказа и конец подзапроса\*/

SELECT

"month", /\*столбец, содержащий данные по месяцам из подзапроса\*/

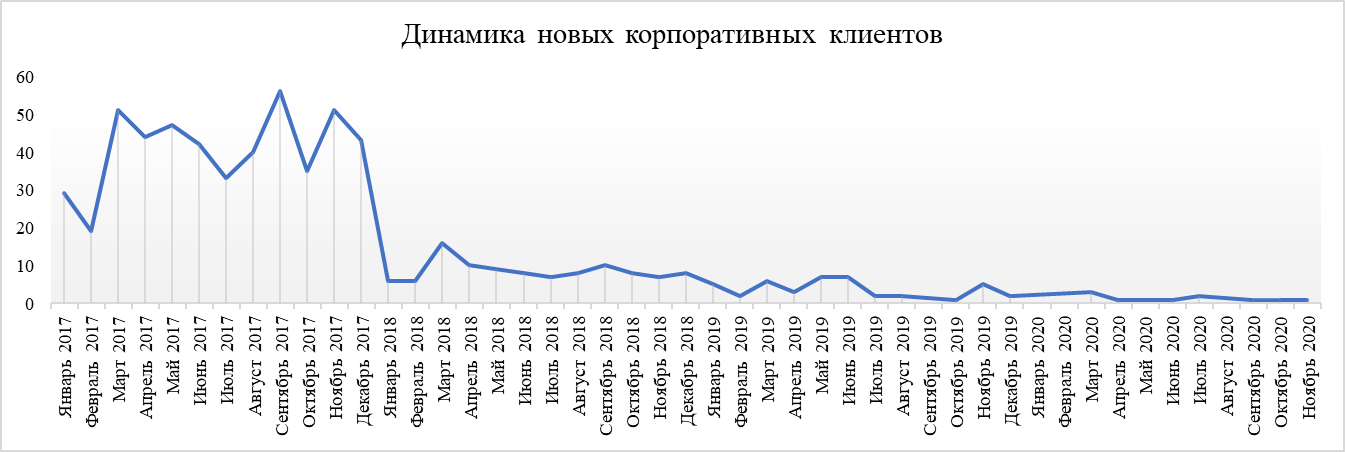
SUM(cust\_cnt) cust\_cnt /\*столбец, содержащий суммарное количество корпоративных клиентов из подзапроса за указанный месяц через функцию SUM\*/

FROM new\_cust /\*из подзапроса new\_cust\*/

GROUP BY 1 /\*группировка данных по месяцу\*/

ORDER BY 1 /\*сортировка данных месяцам\*/

На основании результатов запроса был построен график, отражающий динамику появления новых корпоративных клиентов в компании.



Можно сделать вывод о сильном падении числа новых корпоративных клиентов на протяжении длительного времени. Необходимо продумать новую рекламную стратегию для привлечения новых клиентов.

## 2.3. Составление характеристики В2В клиентов

Для составления характеристики В2В клиента составлен следующий запрос, выводящий средние показатели эффективности по работе с ними:

WITH B2B\_orders AS

(SELECT

s\_cust.cust\_id,

sd.order\_id,

SUM(sc.quantity) quantity,

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount)),1) revenue

FROM

sql.store\_customers s\_cust

LEFT JOIN sql.store\_delivery sd

ON s\_cust.cust\_id=sd.cust\_id

JOIN sql.store\_carts sc

ON sd.order\_id = sc.order\_id

JOIN sql.store\_products sp

ON sc.product\_id=sp.product\_id

GROUP BY s\_cust.cust\_id, sd.order\_id, s\_cust.category

HAVING s\_cust.category = 'Corporate'

ORDER BY 1),

B2B\_address AS

(SELECT

s\_cust.cust\_id,

COUNT (DISTINCT sd.zip\_code) zip\_code

FROM

sql.store\_customers s\_cust

LEFT JOIN sql.store\_delivery sd

ON s\_cust.cust\_id=sd.cust\_id

GROUP BY s\_cust.cust\_id, s\_cust.category

HAVING s\_cust.category = 'Corporate')

SELECT

ROUND (AVG(quantity),1) avg\_quantity,

ROUND (SUM(revenue)/COUNT(revenue),1) avg\_revenue,

ROUND (SUM(zip\_code)/COUNT(zip\_code),1) avg\_addresses

FROM

B2B\_orders

JOIN B2B\_address

ON B2B\_orders.cust\_id = B2B\_address.cust\_id

Детализация и комментарии

WITH B2B\_orders AS /\*ключевое слово WITH и наименование подзапроса\*/

(SELECT /\*начало подзапроса\*/

s\_cust.cust\_id, /\*столбец, содержащий cust\_id\*/

sd.order\_id, /\*столбец, содержащий order\_id\*/

SUM(sc.quantity) quantity, /\*столбец, содержащий сумму товаров в заказе, рассчитаную через функцию SUM\*/

ROUND (SUM((sc.quantity\*sp.price)\*(1-sc.discount)),1) revenue /\*столбец, выводящий выручку по товарам за вычетом скидки, округленной до одного знака после запятой при помощи функции ROUND\*/

FROM

sql.store\_customers s\_cust /\*из таблицы со списком клиентов\*/

JOIN sql.store\_delivery sd /\*соединяем с таблицей со списком доставок\*/

ON s\_cust.cust\_id=sd.cust\_id /\*через cust\_id\*/

JOIN sql.store\_carts sc /\*соединяем через таблицу со списком доставок с таблицей со списком товаров в заказах\*/

ON sd.order\_id = sc.order\_id /\*через order\_id\*/

JOIN sql.store\_products sp /\*соединяем через таблицу со списком товаров в заказах с таблицей со списком продуктов\*/

ON sc.product\_id=sp.product\_id /\*через product\_id\*/

GROUP BY s\_cust.cust\_id, sd.order\_id, s\_cust.category /\*группировка данных по cust\_id, order\_id и категории товара\*/

HAVING s\_cust.category = 'Corporate' /\*фильтрация данных по категории клиента равной ‘Corporate’\*/

ORDER BY 1), /\*сортировка данных по cust\_id и конец подзапроса\*/

B2B\_address AS /\*наименование подзапроса\*/

(SELECT /\*начало подзапроса\*/

s\_cust.cust\_id, /\*столбец, содержащий cust\_id\*/

COUNT (DISTINCT sd.zip\_code) zip\_code /\*столбец, содержащий количество уникальных индексов доставки клиентов\*/

FROM

sql.store\_customers s\_cust /\*из таблицы со списком клиентов\*/

LEFT JOIN sql.store\_delivery sd /\*соединяем слева с таблицей со списком доставок\*/

ON s\_cust.cust\_id=sd.cust\_id /\*через cust\_id\*/

GROUP BY s\_cust.cust\_id, s\_cust.category /\*группировка данных по cust\_id, order\_id и категории товара\*/

HAVING s\_cust.category = 'Corporate') /\*фильтрация данных по категории клиента равное ‘Corporate’ и конец подзапроса\*/

SELECT

ROUND (AVG(quantity),1) avg\_quantity, /\*столбец, содержащий среднее количество товаров клиента, полученный при помощи функции AVG, округленный до одного знака после запятой функцией ROUND\*/

ROUND (SUM(revenue)/COUNT(revenue),1) avg\_revenue, /\*столбец, содержащий среднюю выручку клиента, округленный до одного знака после запятой функцией ROUND, рассчитан следующим образом: сумма всех заказов через функцию SUM поделена на количество всех заказов через функцию COUNT всех строк, содержащих данные о выручке в подзапросе В2В\_orders \*/

ROUND (SUM(zip\_code)/COUNT(zip\_code),1) avg\_addresses /\*столбец, содержащий среднее количество адресов клиента, округленный до одного знака после запятой функцией ROUND, рассчитан следующим образом: деление суммы адресов каждого клиента через функцию SUM на количество клиентов (выраженном в количестве строк содержащих zip\_code в подзапросе) через функцию COUNT\*/

FROM

B2B\_orders /\*из подзапроса В2В\_orders\*/

JOIN B2B\_address /\*соединяем с подзапросом В2В\_address\*/

ON B2B\_orders.cust\_id = B2B\_address.cust\_id /\*через cust\_id\*/

В результате получены следующие характеристики корпоративного клиента:

* среднее количество товаров в заказе у корпоративного клиента - 7,6;
* средняя выручка, получаемая с корпоративного клиента - 285,9;
* среднее количество различных адресов доставки на одного корпоративного клиента - 7,2.

# Анализ логистики компании

## 3.1. Определение эффективности логистики

Для определения эффективности логистики компании необходимо вычислить, какое количество доставок осуществляется вовремя по каждому типу доставки. Для этого составлен следующий запрос:

WITH delivery\_time AS

(SELECT

ship\_mode,

order\_id,

CASE WHEN ship\_mode = 'Standard Class' AND (ship\_date-order\_date) < 7 THEN 'on time'

WHEN ship\_mode ='Second Class' AND (ship\_date-order\_date) < 5 THEN 'on time'

WHEN ship\_mode ='First Class' AND (ship\_date-order\_date) < 4 THEN 'on time'

WHEN ship\_mode ='Same Day' AND (ship\_date-order\_date) < 1 THEN 'on time'

ELSE 'not on time' END timing

FROM

sql.store\_delivery)

SELECT

ship\_mode,

COUNT(order\_id) orders\_cnt,

COUNT(CASE WHEN timing = 'not on time' THEN order\_id END) late\_orders\_cnt,

ROUND (COUNT(CASE WHEN timing = 'on time' THEN order\_id END)::numeric / COUNT(order\_id)::numeric \* 100, 2) "% success"

FROM

delivery\_time

GROUP BY 1

ORDER BY 4

Детализация и комментарии:

WITH delivery\_time AS /\*ключевое слово WITH и наименование подзапроса\*/

(SELECT /\*начало подзапроса\*/

ship\_mode, /\*столбец, содержащий название типа доставки\*/

order\_id, /\*столбец, содержащий order\_id\*/

CASE WHEN ship\_mode = 'Standard Class' AND (ship\_date-order\_date) < 7 THEN 'on time' /\*оператор условия; условие: запись столбца ship\_mode является 'Standard Class' и разница между датой доставки и датой заказа меньше 7, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='Second Class' AND (ship\_date-order\_date) < 5 THEN 'on time' /\*условие: запись столбца ship\_mode является 'Second Class' и разница между датой доставки и датой заказа меньше 5, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='First Class' AND (ship\_date-order\_date) < 4 THEN 'on time' /\*условие: запись столбца ship\_mode является 'First Class' и разница между датой доставки и датой заказа меньше 4, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='Same Day' AND (ship\_date-order\_date) < 1 THEN 'on time' /\*условие: запись столбца ship\_mode является 'Same Day' и разница между датой доставки и датой заказа меньше 1, если так, ставим результат 'on time'\*/

ELSE 'not on time' END timing /\* иначе ставим результат 'not on time' конец действия оператора условия, алиас timing\*/

FROM

sql.store\_delivery) /\*из таблицы со списком доставок и конец подзапроса\*/

SELECT

ship\_mode, /\*столбец, содержащий название типа доставки\*/

COUNT(order\_id) orders\_cnt, /\*столбец, содержащий количество заказов соответствующего типа\*/

COUNT(CASE WHEN timing = 'not on time' THEN order\_id END) late\_orders\_cnt,/\*столбец, содержащий процент заказов, доставленных не вовремя, соответствующих условию, оператор условия: если timing соответствует 'not on time', тогда считаем order\_id через функцию COUNT и конец действия оператора условия\*/

ROUND (COUNT(CASE WHEN timing = 'on time' THEN order\_id END)::numeric / COUNT(order\_id)::numeric \* 100, 2) "% success" /\*столбец, содержащий соотношение количества заказов, которые соответствуют условиям, оператор условия: если timing соответствует 'on time', тогда считаем order\_id через функцию COUNT и конец оператора условия, приведенного в тип данных numeric c общим количеством заказов, посчитанных оператором COUNT, приведенных в тип данных numeric умноженных на 100, чтобы получить число процентов, округленное до двух знаков после запятой при помощи функции ROUND \*/

FROM

delivery\_time /\*из подзапроса delivery\_time\*/

GROUP BY 1 /\*группировка данных по типу доставки\*/

ORDER BY 4 /\*сортировка данных по проценту заказов, доставленных не вовремя\*/

Результат выполнения данного запроса представлен в следующей таблице:

| **ship\_mode** | **orders\_cnt** | **late\_orders\_cnt** | **% success** |
| --- | --- | --- | --- |
| Second Class | 964 | 202 | 79,05 |
| Standard Class | 2 994 | 309 | 89,68 |
| Same Day | 264 | 12 | 95,45 |
| First Class | 787 | 1 | 99,87 |

Можно сделать вывод о том, что чаще всего не вовремя доставляют доставки второго класса.

## 3.2. Определение динамики несвоевременных доставок второго класса

Для определения динамики несвоевременных доставок второго класса составлен следующий запрос:

WITH delivery\_time AS

(SELECT

ship\_mode,

order\_id,

DATE\_TRUNC ('QUARTER', order\_date)::date date,

CASE WHEN ship\_mode = 'Standard Class' AND (ship\_date-order\_date) < 7 THEN 'on time'

WHEN ship\_mode ='Second Class' AND (ship\_date-order\_date) < 5 THEN 'on time'

WHEN ship\_mode ='First Class' AND (ship\_date-order\_date) < 4 THEN 'on time'

WHEN ship\_mode ='Same Day' AND (ship\_date-order\_date) < 1 THEN 'on time'

ELSE 'not on time' END timing

FROM

sql.store\_delivery)

SELECT

date,

ROUND (COUNT(CASE WHEN timing = 'not on time' THEN order\_id END)::numeric / COUNT(order\_id)::numeric \* 100, 2) "% late"

FROM

delivery\_time

GROUP BY date, ship\_mode

HAVING ship\_mode = 'Second Class'

Детализация и комментарии:

WITH delivery\_time AS /\*ключевое слово WITH и наименование подзапроса\*/

(SELECT /\*начало подзапроса\*/

ship\_mode, /\*столбец, содержащий название типа доставки\*/

order\_id, /\*столбец, содержащий order\_id\*/

DATE\_TRUNC ('QUARTER', order\_date)::date date, /\*столбец, выводящий дату заказа, округленную до квартала при помощи функции DATE\_TRUNC и приведенный в тип данных data\*/

CASE WHEN ship\_mode = 'Standard Class' AND (ship\_date-order\_date) < 7 THEN 'on time' /\*оператор условия; условие: запись столбца ship\_mode является 'Standard Class' и разница между датой доставки и датой заказа меньше 7, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='Second Class' AND (ship\_date-order\_date) < 5 THEN 'on time' /\*условие: запись столбца ship\_mode является 'Second Class' и разница между датой доставки и датой заказа меньше 5, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='First Class' AND (ship\_date-order\_date) < 4 THEN 'on time' /\*условие: запись столбца ship\_mode является 'First Class' и разница между датой доставки и датой заказа меньше 4, если так, ставим результат 'on time'\*/

WHEN ship\_mode ='Same Day' AND (ship\_date-order\_date) < 1 THEN 'on time' /\*условие: запись столбца ship\_mode является 'Same Day' и разница между датой доставки и датой заказа меньше 1, если так, ставим результат 'on time'\*/

ELSE 'not on time' END timing /\* иначе ставим результат 'not on time' конец действия оператора условия, алиас timing\*/

FROM

sql.store\_delivery) /\*из таблицы со списком доставок и конец подзапроса\*/

SELECT

date, /\*столбец, содержащий даты, округленные до кварталов\*/

ROUND (COUNT(CASE WHEN timing = 'not on time' THEN order\_id END)::numeric / COUNT(order\_id)::numeric \* 100, 2) "% late" /\*столбец, содержащий соотношение количества доставок, совершенных не вовремя при помощи условия: соответствующие условию timing соответствует ‘not on time’ приведенных к типу numeric с общим количеством доставок, приведенные к типу данных numeric, умноженных на 100 для получения количества процентов, округленных до двух знаков после запятой при помощи функции ROUND\*/

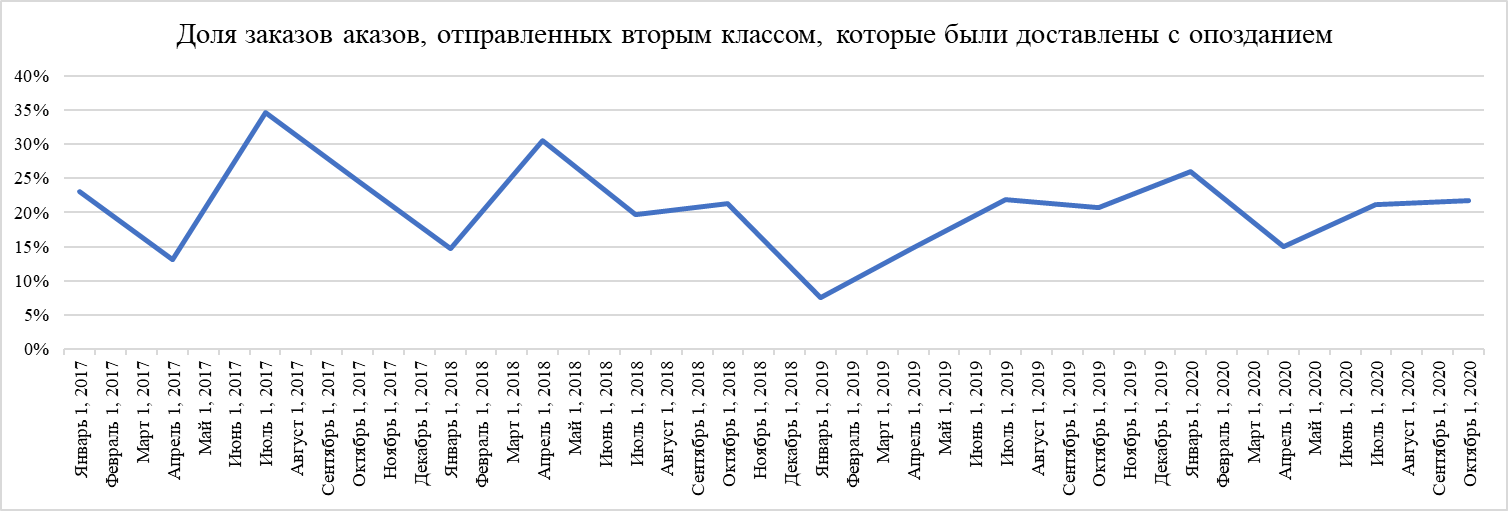
FROM

delivery\_time /\*из подзапроса delivery\_time\*/

GROUP BY date, ship\_mode /\*группировка данных по дате и типу доставки\*/

HAVING ship\_mode = 'Second Class' /\*фильтрация данных типу доставки ‘Second Class’\*/

По результатам запроса построен следующий график:



По данному графику можно сделать вывод о том, что чаще всего проблемы с доставкой возникают в июле, а меньше всего опозданий совершают в январе и апреле. Кажется целесообразным подключение дополнительных мощностей (например, увеличения штата курьеров) в июле для минимизации возможности опозданий в этот период.

## 3.2. Определение штата, который является наиболее перспективным для открытия оффлайн-магазина

В первую очередь необходимо проанализировать в какие штаты осуществляется наибольшее число доставок. Для этого составлен следующий запрос:

SELECT

state,

COUNT(order\_id)

FROM

sql.store\_delivery

GROUP BY 1

ORDER BY 2 desc

Детализация и комментарии:

SELECT

state, /\*столбец, содержащий название штата\*/

COUNT(order\_id) /\*столбец, содержащий количество заказов, осуществленных в конкретный штат при помощи функции COUNT\*/

FROM

sql.store\_delivery /\*из таблицы, содержащий список доставок\*/

GROUP BY 1 /\*группировка данных по штату\*/

ORDER BY 2 desc /\*сортировка по количеству доставок в обратном порядке\*/

По результатам запроса выявлено 5 наиболее популярных штатов доставок, где открытие оффлайн магазина будет перспективным:

* California - 1,021 доставки;
* New York - 562 доставки;
* Texas - 487 доставок;
* Pennsylvania - 288 доставок;
* Illinois - 276 доставок.

Далее, необходимо выделить наиболее популярные города, куда осуществляется доставка. Для этого составлен следующий запрос:

SELECT

state,

city,

COUNT(order\_id)

FROM

sql.store\_delivery

GROUP BY 1, 2

ORDER BY 3 desc

Детализация и комментарии:

SELECT

state, /\*столбец, содержащий название штата\*/

city, /\*столбец, содержащий название города\*/

COUNT(order\_id) /\*столбец, содержащий количество заказов, осуществленных в конкретный штат при помощи функции COUNT\*/

FROM

sql.store\_delivery /\*из таблицы, содержащий список доставок\*/

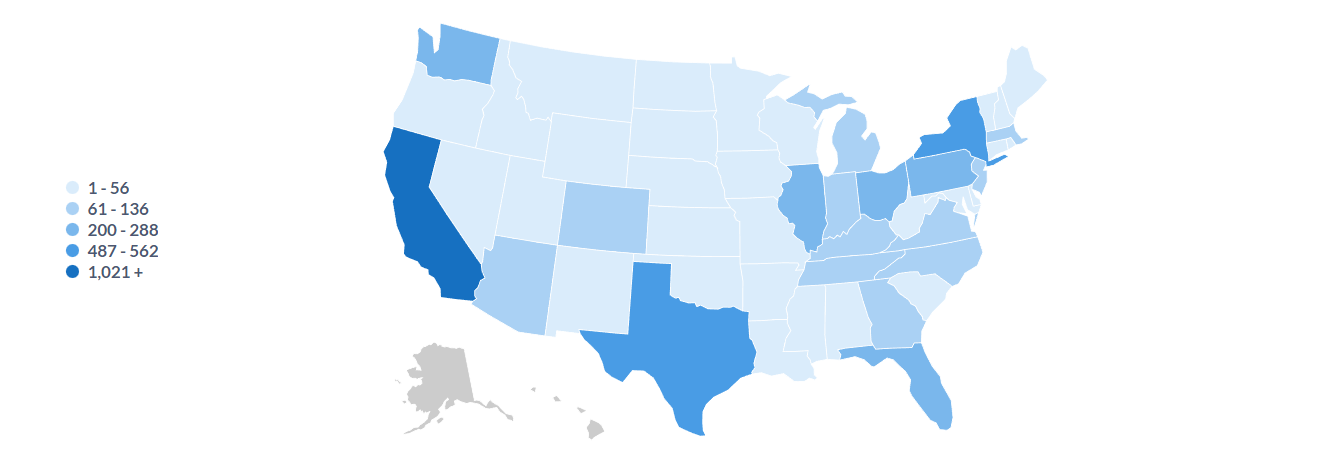
GROUP BY 1, 2 /\*группировка данных по штату и городу\*/

ORDER BY 3 desc /\*сортировка по количеству доставок в обратном порядке\*/

По результатам запроса составлена следующая таблица:

| **state** | **city** | **count** |
| --- | --- | --- |
| New York | New York City | 450 |
| California | Los Angeles | 384 |
| California | San Francisco | 265 |
| Pennsylvania | Philadelphia | 265 |
| Washington | Seattle | 212 |

Можно сделать вывод, что несмотря на то, что больше всего доставок осуществляется в город Нью-Йорк, Лос-Анджелес и Сан-Франциско в сумме превышают его, следовательно, показатели штата Калифорния выше.

Для более наглядной визуализации составлена карта: 

Рекомендованный штат для открытия оффлайн-магазина - Калифорния. В случае, если открытие магазина в Калифорнии невозможно, то следующим подходящим штатом является Калифорния.

Подводя итог можно сделать следующие выводы:

1. Общий объем продаж компании существенно вырос по сравнению с началом работы, хотя в декабре 2020 года наблюдается падение в сравнении с ноябрем. Можно выделить сезонность продаж: несколько лет подряд они растут в марте, сентябре и ноябре и падают в январе, феврале и октябре. Рост в сентябре может быть связан с началом офисного сезона, а в ноябре со скидками к “черной пятнице”, рост в марте скорее всего связан с некоторыми особенностями американского рынка. Для снижения падения в январе, феврале и октябре рекомендуется ввести дополнительные скидки и усилить рекламную активность.
2. Наиболее продаваемыми являются следующие категории товаров: офисные кресла, телефоны и органайзеры. Наибольшую выручку приносят следующие наименования: копировальный аппарат Canon imageCLASS 2200 Advanced Copier, офисное кресло HON 5400 Series Task Chairs for Big and Tall и принтер 3D Systems Cube Printer, 2nd Generation, Magenta.
3. Наибольшую выручку компании приносят корпоративные клиенты. До декабря 2017 года число новых корпоративных клиентов было существенным, а далее последовало стремительное падение этого показателя и их количество не превышало 10 в месяц. Необходимо проработать рекламную стратегию компании для привлечения большего числа новых корпоративных клиентов. В среднем корпоративные клиенты заказывают 8 различных товаров в 1 заказе, приносят среднюю выручку в 285,9 долларов, и имеют 7 адресов доставки.
4. Около 10 процентов всех доставок компания осуществляет с опозданиями, выше всего процент опозданий в доставках второго класса. Чаще всего проблемы с доставкой второго класса возникают в июле, а меньше всего опозданий совершают в январе и апреле. Рекомендуется увеличения штата курьеров в июле для минимизации возможности опозданий в этот период.
5. Наибольшее число доставок осуществляется в следующие штаты: Калифорния, Нью-Йорк и Техас. А самыми популярными городами доставок являются Нью-Йорк, Лос-Анджелес и Сан-Франциско. Для открытия оффлайн-магазина приоритетным штатом является Калифорния, а если открытие магазина там невозможно, то Нью-Йорк.