

Оценка эффективности бизнеса компании *Pens and Pencils*

ЗАДАЧА

1. Динамика продаж и распределение выручки по товарам.

Динамика продаж по месяцам

Категории и подкатегории товаров, приносящих наибольший доход

Топ-25 товаров по выручке

2. Портрет клиента

Соотношение B2B и B2C клиентов

Динамика новых корпоративных клиентов

Сколько в среднем различных товаров в заказах у корпоративных клиентов?

Какая в среднем сумма заказов у корпоративных клиентов?

3. Логистика компании

Эффективность доставок

Количество доставок в зависимости от штата

ЗАДАЧА

Проанализировать работу компании с точки зрения её эффективности и дать рекомендации по масштабированию бизнеса, а именно в каком штате лучше открыть офлайн-магазин.

КОНКРЕТНЫЕ ШАГИ

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)

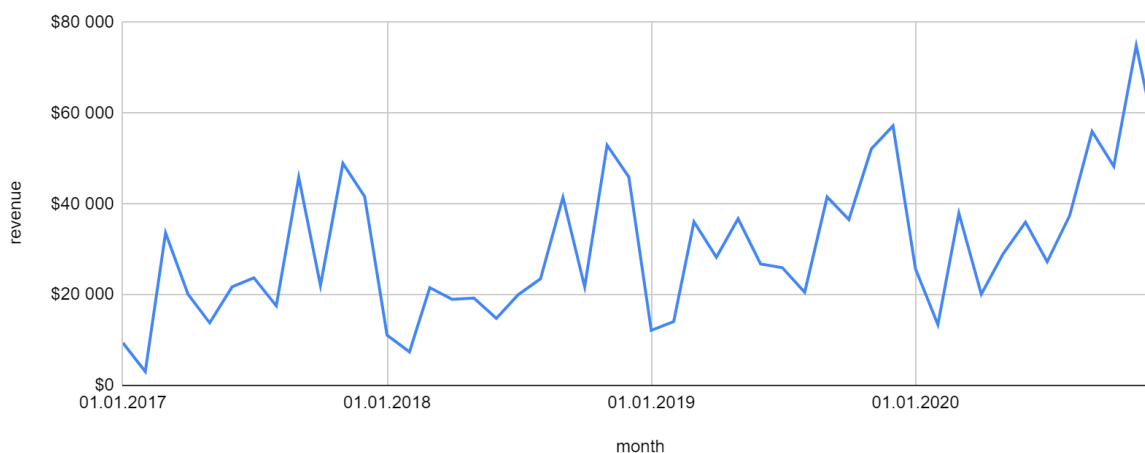
1. Динамика продаж и распределение выручки по товарам.

- Какова **динамика продаж**?
- Какие **категории и подкатегории** продаж приносят больше всего **выручки**?
- Какие **товары** продаются лучше всего?

Динамика продаж по месяцам

```
SELECT
  DATE_TRUNC('month', order_date)::date orders_date, -- месяц заказов
  ROUND(SUM((price - price*discount)*quantity)) revenue -- выручка за месяц
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 DATE_TRUNC('month', order_date)
ORDER BY orders_date
```

Динамика выручки по месяцам



На графике можно заметить, что доходы увеличиваются в конце года и падают в начале каждого года.

Конкретные данные по доходам в каждом месяце представлены ниже:

month	revenue
01.01.2017	\$9 394
01.02.2017	\$3 040
01.03.2017	\$33 638
01.04.2017	\$20 108
01.05.2017	\$13 784
01.06.2017	\$21 739
01.07.2017	\$23 720
01.08.2017	\$17 536
01.09.2017	\$45 852
01.10.2017	\$22 021
01.11.2017	\$48 931
01.12.2017	\$41 632
01.01.2018	\$11 088
01.02.2018	\$7 400
01.03.2018	\$21 547
01.04.2018	\$18 971
01.05.2018	\$19 253
01.06.2018	\$14 764
01.07.2018	\$19 969
01.08.2018	\$23 499
01.09.2018	\$41 462
01.10.2018	\$21 726
01.11.2018	\$52 948
01.12.2018	\$45 949
01.01.2019	\$12 141
01.02.2019	\$14 067
01.03.2019	\$36 057
01.04.2019	\$28 246
01.05.2019	\$36 745
01.06.2019	\$26 777
01.07.2019	\$25 959
01.08.2019	\$20 535
01.09.2019	\$41 531
01.10.2019	\$36 558

01.11.2019	\$52 114
01.12.2019	\$57 208
01.01.2020	\$25 729
01.02.2020	\$13 416
01.03.2020	\$37 962
01.04.2020	\$20 114
01.05.2020	\$28 979
01.06.2020	\$35 985
01.07.2020	\$27 247
01.08.2020	\$37 409
01.09.2020	\$55 987
01.10.2020	\$48 273
01.11.2020	\$74 906
01.12.2020	\$52 241

Категории и подкатегории товаров, приносящих наибольший доход

```

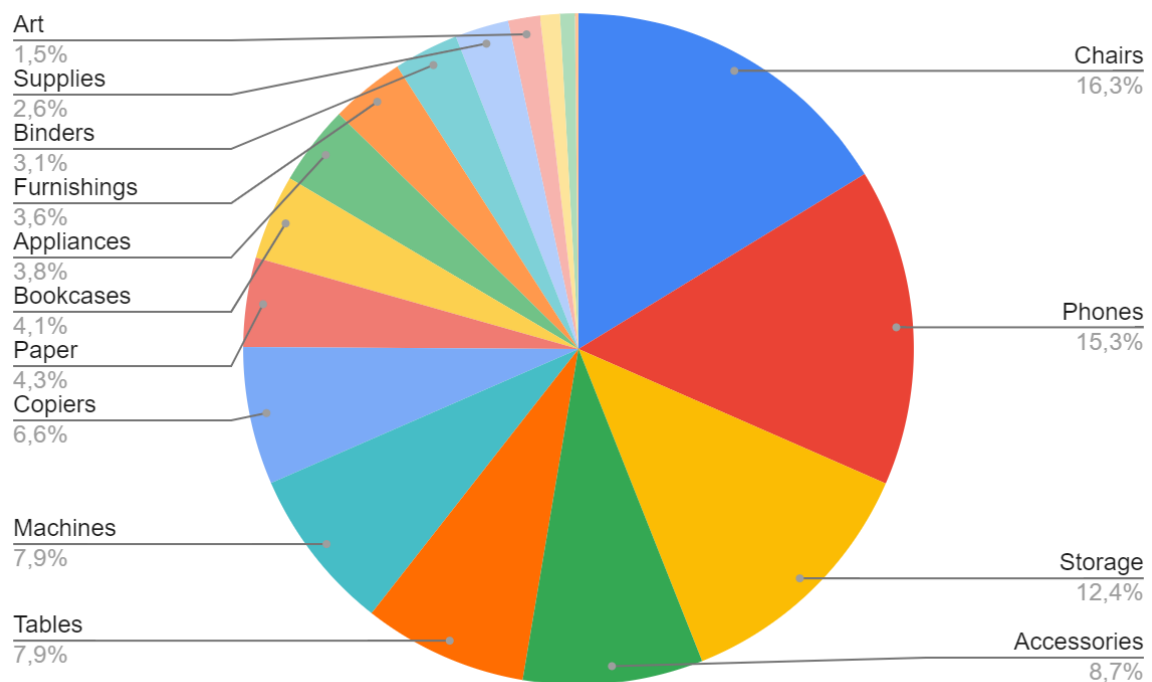
SELECT
    category,
    subcategory,
    ROUND(SUM((price - price*discount)*quantity)) revenue
FROM
    sql.store_carts sc
    JOIN sql.store_products sp ON sc.product_id = sp.product_id
GROUP BY category, subcategory
ORDER BY revenue DESC

```

category	subcategory	revenue
Furniture	Chairs	\$235 318
Technology	Phones	\$221 110
Office Supplies	Storage	\$179 736
Technology	Accessories	\$125 442
Furniture	Tables	\$114 532
Technology	Machines	\$113 978

Technology	Copiers	\$95 844
Office Supplies	Paper	\$62 148
Furniture	Bookcases	\$59 271
Office Supplies	Appliances	\$55 550
Furniture	Furnishings	\$51 695
Office Supplies	Binders	\$45 312
Office Supplies	Supplies	\$37 457
Office Supplies	Art	\$22 148
Office Supplies	Envelopes	\$13 809
Office Supplies	Labels	\$10 345
Office Supplies	Fasteners	\$2 459

На круговой диаграмме можно увидеть процентное соотношение разных групп товаров:



Топ-25 товаров по выручке

```

WITH total_revenue_CTE AS
(
SELECT
    SUM((price - price*discount)*quantity) total_revenue -- общая выручка
FROM
    sql.store_carts sc
    JOIN sql.store_products sp ON sc.product_id = sp.product_id
)

SELECT
    product_nm,
    ROUND(SUM((price - price*discount)*quantity), 2) revenue,
/*выручка по конкр.продукту*/
    SUM(quantity) quantity, -- количество проданных товаров
    ROUND(SUM((price - price*discount)*quantity)/total_revenue*100, 2) percent_from_total
/*процент от общей выручки*/
FROM
    sql.store_carts sc
    JOIN sql.store_products sp ON sc.product_id = sp.product_id
    CROSS JOIN total_revenue_CTE
GROUP BY product_nm, total_revenue
ORDER BY revenue DESC
LIMIT 25

```

product_nm	revenue	quantity	percent_from_total
Canon imageCLASS 2200 Advanced Copier	\$36 960	20	2,56%
HON 5400 Series Task Chairs for Big and Tall	\$15 310	39	1,06%
3D Systems Cube Printer, 2nd Generation, Magenta	\$14 300	11	0,99%
High Speed Automatic Electric Letter Opener	\$13 624	11	0,94%
Martin Yale Chadless Opener Electric Letter Opener	\$13 325	22	0,92%
Cisco TelePresence System EX90 Videoconferencing Unit	\$11 319	6	0,78%
Hewlett Packard LaserJet 3310 Copier	\$11 304	38	0,78%
Samsung Galaxy Mega 6.3	\$11 155	39	0,77%
Canon PC1060 Personal Laser Copier	\$9 296	19	0,64%
HP Designjet T520 Inkjet Large Format Printer - 24" Color	\$9 188	12	0,64%
Global Troy Executive Leather Low-Back Tilter	\$9 083	32	0,63%

Honeywell Enviracaire Portable HEPA Air Cleaner for 17' x 22' Room	\$9 044	40	0,63%
Tennsco 6- and 18-Compartment Lockers	\$8 740	43	0,60%
Plantronics CS510 - Over-the-Head monaural Wireless Headset System	\$8 658	35	0,60%
SAFCO Arco Folding Chair	\$8 101	53	0,56%
Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish	\$7 805	24	0,54%
Apple iPhone 5	\$7 798	24	0,54%
Tennsco Double-Tier Lockers	\$7 705	47	0,53%
Ativa V4110MDD Micro-Cut Shredder	\$7 700	11	0,53%
Hon Deluxe Fabric Upholstered Stacking Chairs, Rounded Back	\$7 446	51	0,51%
Tennsco Single-Tier Lockers	\$7 327	28	0,51%
Office Star - Professional Matrix Back Chair with 2-to-1 Synchro Tilt and Mesh Fabric Seat	\$7 020	28	0,49%
Zebra ZM400 Thermal Label Printer	\$6 966	6	0,48%
Hot File 7-Pocket, Floor Stand	\$6 511	49	0,45%
Bretford Rectangular Conference Table Tops	\$6 498	46	0,45%

В топ-25 товаров представлены товары категорий **Technology, Furniture и Office Supplies**.

2. Портрет клиента

- Как соотносится **количество клиентов** по категориям **B2B** и **B2C**?
- Как соотносится **выручка** по этим категориям?
- Каких клиентов больше, **новых** или **старых**?

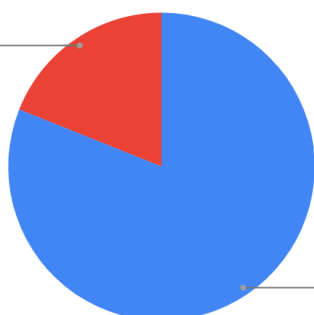
Соотношение B2B и B2C клиентов

```
SELECT
  sc.category, -- категория клиента
  COUNT(DISTINCT sc.cust_id) cust_cnt, -- количество клиентов
  ROUND(SUM((price - price*discount)*quantity)) revenue -- выручка
FROM
  sql.store_customers sc
  JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id
  JOIN sql.store_carts o ON sd.order_id = o.order_id
  JOIN sql.store_products sp ON o.product_id = sp.product_id
GROUP BY sc.category
ORDER BY revenue DESC
```

category	cust_cnt	revenue
Corporate	645	\$1 172 009
Consumer	148	\$274 148

Доход

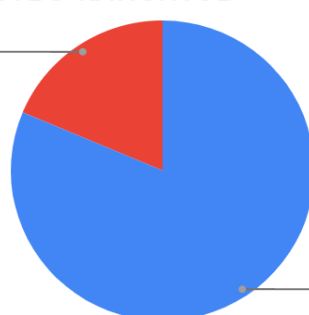
B2C
19,0%



B2B
81,0%

Количество клиентов

B2C
18,7%



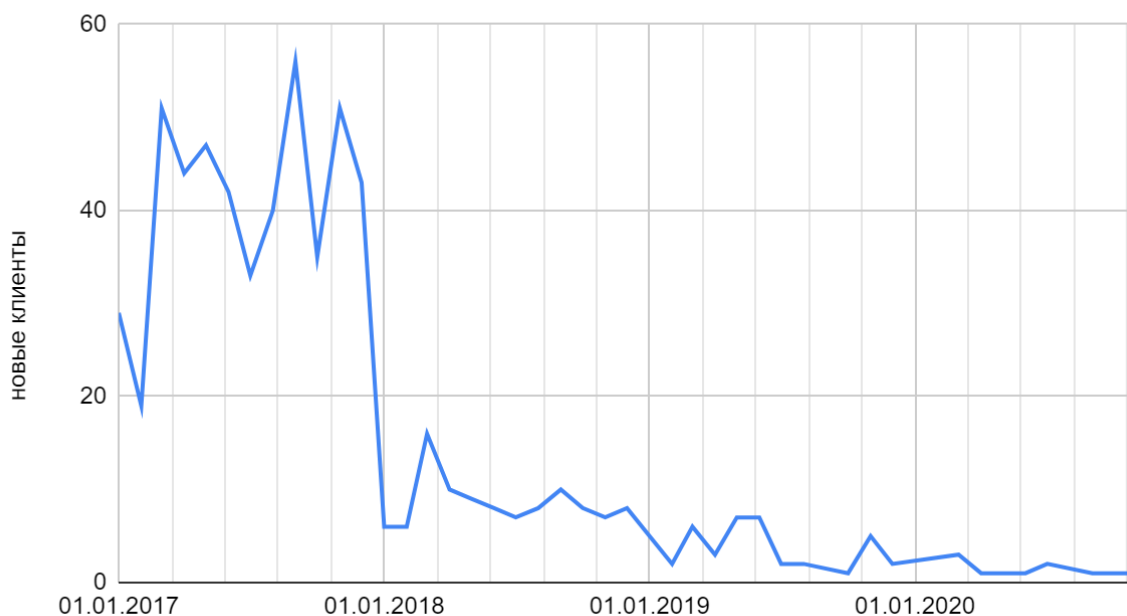
B2B
81,3%

Динамика новых корпоративных клиентов

```
WITH month_cohorts AS
(
SELECT
    sc.cust_id,
    DATE_TRUNC('month', MIN(order_date))::date first_order_month
/*месяц, когда клиентом был сделан первый заказ*/
FROM
    sql.store_customers sc
    JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id
WHERE sc.category = 'Corporate' -- категория B2B
GROUP BY sc.cust_id
)

SELECT
    first_order_month, -- месяц
    COUNT(cust_id) new_custs -- количество новых клиентов
FROM
    month_cohorts
GROUP BY first_order_month
ORDER BY first_order_month
```

Динамика кол-ва новых B2B клиентов



По графику видно, что в конце 2017-го года наблюдается резкое падение количества новых корпоративных клиентов, которое и далее продолжает падать. Следует выявить причины отрицательной динамики по новым клиентам в сегменте B2B. Также мы знаем

из предыдущего раздела, что выручка, тем не менее, все четыре года обнаруживала тенденции к росту. С этой точки зрения привлечение новых корпоративных клиентов представляется особенно важным. Далее представлены данные по каждому месяцу:

first_order_month	new_custs
01.01.2017	29
01.02.2017	19
01.03.2017	51
01.04.2017	44
01.05.2017	47
01.06.2017	42
01.07.2017	33
01.08.2017	40
01.09.2017	56
01.10.2017	35
01.11.2017	51
01.12.2017	43
01.01.2018	6
01.02.2018	6
01.03.2018	16
01.04.2018	10
01.05.2018	9
01.06.2018	8
01.07.2018	7
01.08.2018	8
01.09.2018	10
01.10.2018	8
01.11.2018	7
01.12.2018	8
01.01.2019	5
01.02.2019	2
01.03.2019	6
01.04.2019	3
01.05.2019	7
01.06.2019	7
01.07.2019	2
01.08.2019	2
01.10.2019	1

01.11.2019	5
01.12.2019	2
01.03.2020	3
01.04.2020	1
01.06.2020	1
01.07.2020	2
01.09.2020	1
01.10.2020	1
01.11.2020	1

Сколько в среднем различных товаров в заказах у корпоративных клиентов?

В среднем **в одном заказе** корпоративного клиента **2 различных товара**.

WITH unique_products_per_order_B2B AS

```
(
SELECT
    sc.cust_id, -- id клиента
    sd.order_id, -- id заказа
    COUNT(DISTINCT o.product_id) unique_products_cnt
/*количество различных продуктов*/
FROM
    sql.store_customers sc
    JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id
    JOIN sql.store_carts o ON sd.order_id = o.order_id
WHERE sc.category = 'Corporate'
GROUP BY sc.cust_id, sd.order_id
)
```

```
SELECT
    AVG(unique_products_cnt) /*среднее количество различных продуктов в заказе*/
FROM unique_products_per_order_B2B
```

Какая в среднем сумма заказов у корпоративных клиентов?

Средняя **выручка с одного заказа** у корпоративного клиента составляет **286 \$**.

```
WITH revenue_per_order_B2B AS
(
  SELECT
    o.order_id,
    ROUND(SUM((price - price*discount)*quantity)) order_revenue
  FROM
    sql.store_customers sc
    JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id
    JOIN sql.store_carts o ON sd.order_id = o.order_id
    JOIN sql.store_products sp ON o.product_id = sp.product_id
  WHERE sc.category = 'Corporate'
  GROUP BY o.order_id
)

SELECT
  ROUND(AVG(order_revenue)) -- средняя выручка за 1 заказ
FROM
  revenue_per_order_B2B
```

Сколько в среднем различных офисов у корпоративных клиентов?

На одного корпоративного клиента приходится около **6** различных **офисов**.

```
WITH offices_per_customer_B2B AS
(
  SELECT
    sc.cust_id,
    COUNT(DISTINCT zip_code) office_cnt -- кол-во офисов клиента
  FROM
    sql.store_customers sc
    JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id

  WHERE sc.category = 'Corporate'
  GROUP BY sc.cust_id
)

SELECT
  ROUND(AVG(office_cnt)) -- среднее кол-во офисов на одного B2B клиента
FROM
  offices_per_customer_B2B
```

3. Логистика компании

- Насколько **эффективно** выполняются текущие доставки?
- Как распределяются доставки и выручка **по штатам и городам**?

Эффективность доставок

Оценка эффективности в зависимости от типа доставки:

```
WITH delivery_success AS
(
SELECT
  order_id,
  CASE WHEN ship_mode = 'Standard Class' and ship_date - order_date <= 6 THEN 1
        WHEN ship_mode = 'Second Class' and ship_date - order_date <= 4 THEN 1
        WHEN ship_mode = 'First Class' and ship_date - order_date <= 3 THEN 1
        WHEN ship_mode = 'Same Day' and ship_date - order_date <= 0 THEN 1
        ELSE NULL END success
  /*подсчёт заказов, доставленных в срок, в соответствии с типом доставки*/
FROM sql.store_delivery
)

SELECT
  ship_mode, -- тип доставки
  COUNT(sd.order_id) orders_cnt, -- общее количество заказов
  COUNT(sd.order_id) - SUM(success) late_orders_cnt,
  /*количество заказов, доставленных с опозданием*/
  ROUND(COUNT(success)::decimal/COUNT(sd.order_id)::decimal * 100, 2) "% success"
  /*процент заказов, доставленных в срок*/
FROM
  sql.store_delivery sd
  JOIN delivery_success ds ON sd.order_id = ds.order_id
GROUP BY ship_mode
ORDER BY "% success"
```

ship_mode	orders_cnt	late_orders_cnt	% success
Second Class	964	202	79,05%
Standard Class	2994	309	89,68%
Same Day	264	12	95,45%
First Class	787	1	99,87%

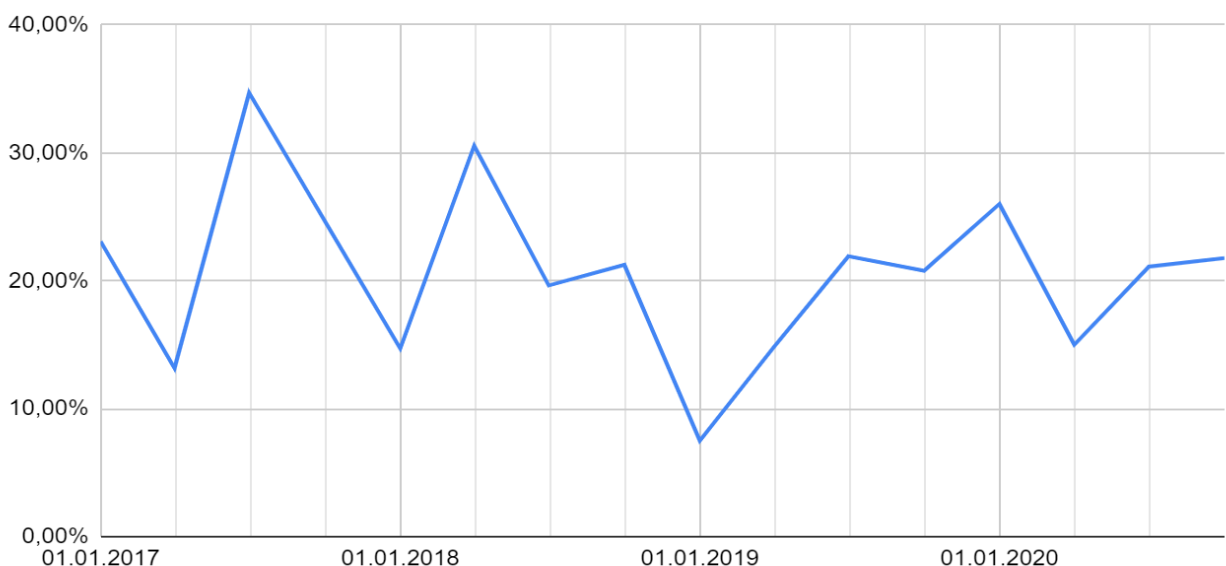
Из таблицы можно сделать вывод, что больше всего **проблем** существует с **доставкой типа Second Class** (только 79% заказов были доставлены вовремя).

Оценка доли доставленных с опозданием заказов с типом доставки Second Class в динамике по кварталам:

```
WITH delivery_success AS
(
SELECT
  order_id,
  CASE WHEN ship_mode = 'Standard Class' and ship_date - order_date <= 6 THEN 1
        WHEN ship_mode = 'Second Class' and ship_date - order_date <= 4 THEN 1
        WHEN ship_mode = 'First Class' and ship_date - order_date <= 3 THEN 1
        WHEN ship_mode = 'Same Day' and ship_date - order_date <= 0 THEN 1
        ELSE NULL END success -- подсчёт доставленных вовремя заказов
FROM sql.store_delivery
)

SELECT
  DATE_TRUNC('quarter', order_date), -- квартал года
  ROUND((COUNT(sd.order_id) - SUM(success))::decimal / COUNT(sd.order_id)::decimal
* 100, 2) late_orders_rate /*процент заказов, доставленных с опозданием*/
FROM
  sql.store_delivery sd
  JOIN delivery_success ds ON sd.order_id = ds.order_id
WHERE ship_mode = 'Second Class' -- тип доставки только Second Class
GROUP BY DATE_TRUNC('quarter', order_date)
ORDER BY DATE_TRUNC('quarter', order_date)
```

Late orders rate per quarter (ship_mode = 'Second Class')



quarter	late_orders_rate
01.01.2017	23,08%
01.04.2017	13,16%
01.07.2017	34,69%
01.10.2017	24,68%
01.01.2018	14,71%
01.04.2018	30,56%
01.07.2018	19,64%
01.10.2018	21,25%
01.01.2019	7,50%
01.04.2019	14,81%
01.07.2019	21,92%
01.10.2019	20,78%
01.01.2020	26,00%
01.04.2020	15,00%
01.07.2020	21,11%
01.10.2020	21,77%

Исходя из графика и данных, доля поздних доставок распределяется по кварталам года относительно равномерно, из чего можно сделать вывод, что доставки второго класса постоянно осуществляются с опозданием.

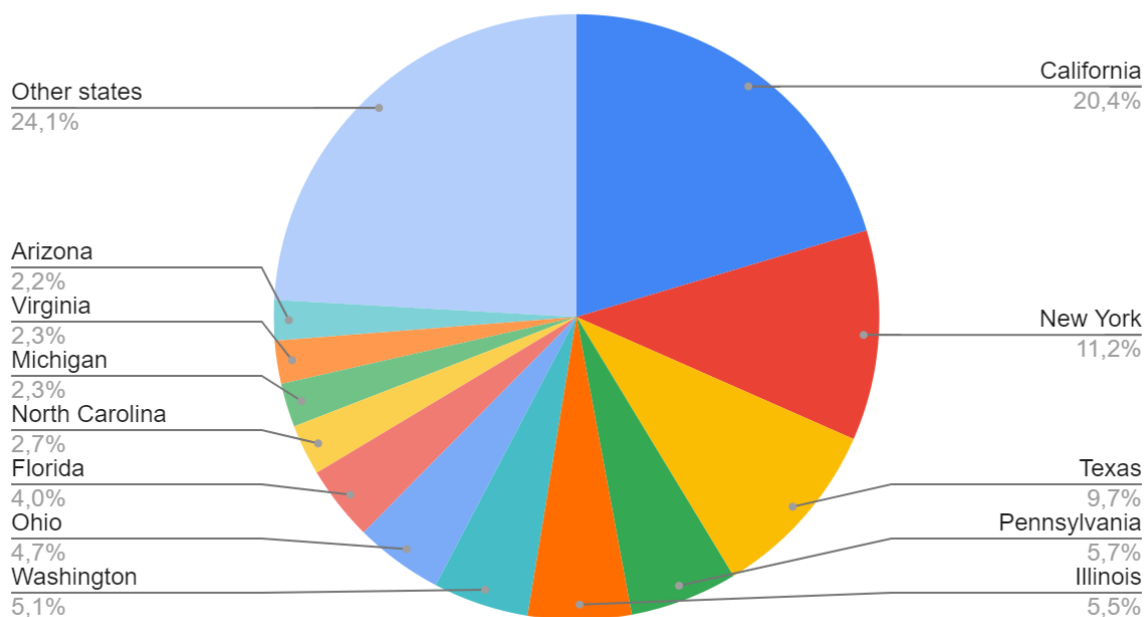
Количество доставок в зависимости от штата

Количество доставок будем считать по количеству заказов на основе таблицы store_delivery.

```
SELECT
    state, -- штат
    COUNT(order_id) orders_cnt -- кол-во заказов в штате за весь период
FROM
    sql.store_delivery
GROUP BY state
ORDER BY COUNT(order_id) DESC
```

На представленной ниже диаграмме видно, что самым перспективным штатом для открытия офлайн-магазина является **Калифорния**, т.к. доля заказов в этом штате составляет $\frac{1}{5}$ **всех заказов** (более 20%):

Количество заказов в зависимости от штата (%)



Лидер среди городов по количеству заказов - Нью-Йорк. Тем не менее, отдать предпочтение стоит именно Калифорнии. Мы видим на диаграмме и в приведённой ниже таблице, что поставки в Калифорнию почти вдвое превышают поставки в Нью-Йорк.

Штаты Нью-Йорк и Техас следуют друг за другом с небольшим отрывом, Нью-Йорк - следующий потенциальный кандидат для открытия новой офлайн-площадки при дальнейшем развитии бизнеса.

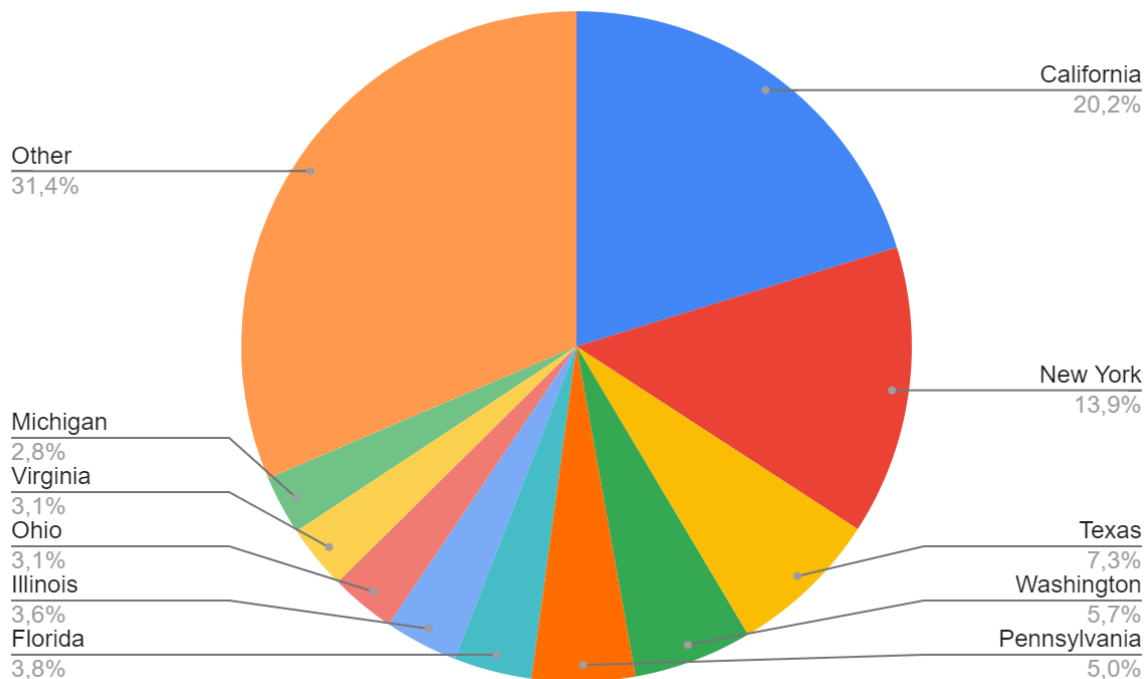
state	orders_cnt
California	1 021
New York	562
Texas	487
Pennsylvania	288
Illinois	276
Washington	256
Ohio	236
Florida	200
North Carolina	136
Michigan	117
Virginia	115
Arizona	108
Georgia	91
Tennessee	91

Colorado	79
Indiana	73
Massachusetts	62
New Jersey	61
Kentucky	61
Oregon	56
Wisconsin	53
Maryland	46
Connecticut	45
Minnesota	44
Delaware	44
Alabama	34
Oklahoma	34
Missouri	30
Arkansas	27
Mississippi	26
Utah	26
Rhode Island	25
Nebraska	23
Nevada	23
New Mexico	22
Louisiana	21
South Carolina	20
New Hampshire	17
Iowa	17
Kansas	14
Idaho	11
Montana	8
Vermont	6
South Dakota	5
District of Columbia	4
Maine	3
West Virginia	2
North Dakota	2
Wyoming	1

Следующий запрос позволяет оценить **сумму выручки в каждом конкретном штате**:

```
SELECT
  sd.state,
  ROUND(SUM((price - price*discount)*quantity)) revenue -- выручка в конкретном штате
FROM
  sql.store_customers sc
  JOIN sql.store_delivery sd ON sc.cust_id = sd.cust_id
  JOIN sql.store_carts o ON sd.order_id = o.order_id
  JOIN sql.store_products sp ON o.product_id = sp.product_id
GROUP BY sd.state
ORDER BY revenue DESC
```

Мы видим, что и по выручке штат Калифорния занимает беспспорное первое место и также приносит $\frac{1}{5}$ **ВСЕГО ДОХОДА** компании.



state	revenue
California	\$292 430
New York	\$201 703
Texas	\$105 377
Washington	\$82 827
Pennsylvania	\$71 776
Florida	\$54 399

Illinois	\$51 343
Ohio	\$45 438
Virginia	\$45 054
Michigan	\$41 133
Indiana	\$34 085
North Carolina	\$32 833
Georgia	\$29 004
New Jersey	\$27 470
Kentucky	\$23 124
Arizona	\$22 904
Wisconsin	\$22 726
Colorado	\$18 853
Massachusetts	\$18 514
Tennessee	\$17 770
Delaware	\$17 490
Missouri	\$16 413
Alabama	\$15 403
Minnesota	\$15 386
Maryland	\$14 809
Rhode Island	\$14 802
Oklahoma	\$12 711
Oregon	\$11 625
Nevada	\$9 222
Connecticut	\$9 022
Utah	\$7 725
Arkansas	\$7 626
Mississippi	\$7 213
South Carolina	\$6 246
Louisiana	\$6 101
Vermont	\$5 239
New Hampshire	\$5 119
Montana	\$4 087
New Mexico	\$3 589
Nebraska	\$3 308
Iowa	\$3 019
Idaho	\$2 731

District of Columbia	\$2 425
Kansas	\$1 924
Wyoming	\$1 122
Maine	\$951
South Dakota	\$930
North Dakota	\$721
West Virginia	\$433

4. Общие выводы

1. С точки зрения **доходов** дела компании в целом идут достаточно хорошо, в течение четырёх лет наблюдается их **повышение**. Доходы компании имеют тенденцию увеличиваться в конце года (осень - начало зимы). **Пик** обычно приходится на **ноябрь**, и если в 2017-м году доход в ноябре составил около 49 тыс. долларов, то в ноябре 2020-го - около 75 тыс. долларов.
2. **Лидеры продаж - кресла/стулья, телефоны и офисные принадлежности.** Возможно, следует выделить средства на рекламу столов, которые занимают пятое место в рейтинге подкатегорий товаров и могли бы идти как удачное дополнение к покупке стульев, которые являются лидерами продаж. В топ-25 товаров представлены товары категорий Technology, Furniture и Office Supplies.
3. Более **80%** как приносимого **дохода**, так и **клиентов** приходится на **сегмент B2B**. Именно инвестиции в эту сферу принесут львиную долю выручки. Один заказ корпоративного клиента - это в среднем 2 товара и около 300 \$. Удержание корпоративных клиентов находится на хорошем уровне - это видно по динамике дохода компании, однако **с конца 2017-го года привлечение новых B2B клиентов** находится на **низком уровне** и продолжает падать - следует выявить причины падения количества новых клиентов и подготовить рекламные кампании, которые могли бы привлечь их.
4. **Доставка** в большинстве случаев осуществляется **вовремя** (это 90% и выше), однако исключением является тип доставки **Second Class**, у которого пятая часть всех доставок были выполнены с **опозданием** - следует выявить возможные существующие проблемы в этой области.
5. Однозначным лидером среди штатов и по количеству доставок, и по общей выручке является **Калифорния** - на этот штат приходится $\frac{1}{5}$ **всех заказов и доходов**. Поэтому рекомендуется открыть офлайн-магазин именно в этом штате.
6. Следует также заметить, что 14% прибыли приходится на Нью-Йорк, поэтому в дальнейшем рекомендовано открыть офлайн-магазин также и в этом штате.