

Цели проекта

1. Оценить динамику продаж и распределение выручки по товарам.
2. Составить портрет клиента, а для этого — 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).

Анализ эффективности продаж

1. Динамика продаж

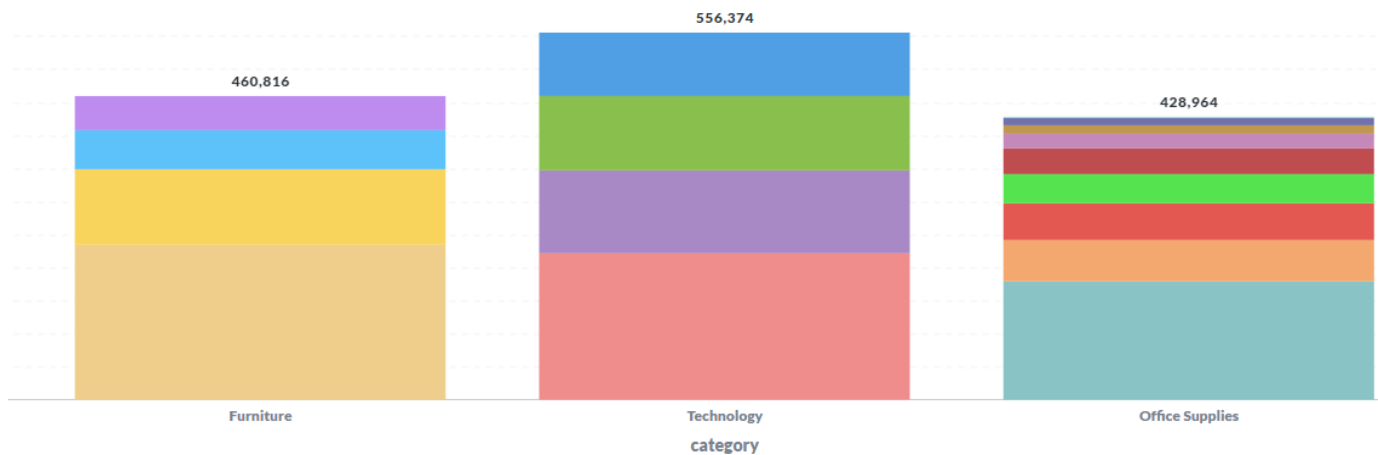
```
SELECT
  date_trunc('month', d.order_date)::date date,
  ROUND(SUM((p.price*c.quantity)*(1-c.discount))) revenue
FROM sql.store_delivery d
JOIN sql.store_carts c USING(order_id)
JOIN sql.store_products p USING(product_id)
GROUP BY 1
ORDER BY 1
```



Динамика выручки за три года выросла, при этом присутствует определенная сезонность: в сентябре, ноябре и декабре наблюдаются скачки продаж.

2. Выручка по категориям и подкатегориям

```
SELECT
  p.category,
  p.subcategory,
  ROUND(SUM((p.price*c.quantity)*(1-c.discount))) revenue
FROM sql.store_delivery d
JOIN sql.store_carts c USING(order_id)
JOIN sql.store_products p USING(product_id)
GROUP BY 1, 2
ORDER BY revenue DESC
```



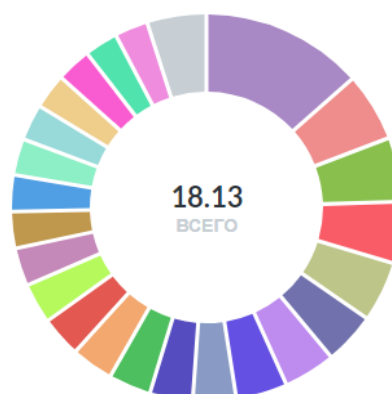
Больше всего выручки за все время принесла категория Technology, а самая доходная подкатегория - Chairs из категории Furniture

3. Топ-25 товаров по объёму выручки и доля, которую принес каждый товар от общей выручки

```
WITH revenue_product AS --CTE отбора товаров, их количества и
выручки
    (SELECT
        DISTINCT (p.product_nm),
        SUM(c.quantity) quantity,
        SUM((c.quantity*p.price)*(1-c.discount)) revenue
    FROM sql.store_products p
    JOIN sql.store_carts c USING(product_id)
    GROUP BY 1),
total_revenue AS --CTE расчета суммарной выручки по всем товарам
    (SELECT
        sum(revenue) AS sum_revenue
    FROM revenue_product)
SELECT
    rp.product_nm,
    rp.revenue,
    rp.quantity,
    (rp.revenue/tr.sum_revenue)*100 AS percent_from_total
FROM revenue_product rp
CROSS JOIN total_revenue tr
GROUP BY 1, 2, 3, 4
ORDER BY 2 DESC
LIMIT 25
```

Наименование товара	Выручка	Количество проданных товаров	Доля от общей выручки, %
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	13 325,00	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	9 187,50	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	0,60
Plantronics CS510 - Over-the-Head monaural Wireless Headset System	8 657,89	35	0,60
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	7 798,00	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	7 019,50	28	0,49
Zebra ZM400 Thermal Label Printer	6 965,70	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

- Canon imageCLASS 2200 Advanced Copier
- HON 5400 Series Task Chairs for Big and Tall
- 3D Systems Cube Printer, 2nd Generation, Magenta
- High Speed Automatic Electric Letter Opener
- Martin Yale Chadless Opener Electric Letter Opener
- Cisco TelePresence System EX90 Videoconferencing Unit
- Hewlett Packard LaserJet 3310 Copier
- Samsung Galaxy Mega 6.3
- Canon PC1060 Personal Laser Copier
- HP Designjet T520 Inkjet Large Format Printer - 24" Color
- Global Troy Executive Leather Low-Back Tilter
- Honeywell Enviracaire Portable HEPA Air Cleaner for 17' x 22' Ro...
- Tennsco 6- and 18-Compartment Lockers
- 11 больше



Топ-25 товаров принесли 18.13 % от всей выручки компании.

4. Вывод

Анализ эффективности продаж показал, что в целом, динамика продаж положительная. В продажах присутствует сезонность: в сентябре, ноябре и декабре скачки продаж, в январе и феврале - провалы продаж.

Самая доходная категория - Technology. Категории Furniture и Office Supplies принесли примерно одинаково дохода.

Самая доходная подкатегория - Chairs, она принесла дохода чуть больше, чем подкатегория Phones.

Топ-25 товаров по объему выручки принесли 18.13 % от всей выручки компании. Это говорит о том, что ассортимент компании достаточно широкий.

Рекомендация: используя знание о сезонности, необходимо увеличивать затраты на маркетинг в несезон, а также в предсезонные месяцы.

Портрет клиента

1. Соотношение количества клиентов и выручки по категориям B2B и B2C

```
SELECT
  cu.category,
  count(DISTINCT d.cust_id) AS cust_cnt,
  ROUND(SUM((p.price*c.quantity)*(1-c.discount))) revenue
FROM sql.store_customers cu
JOIN sql.store_delivery d USING(cust_id)
JOIN sql.store_carts c USING(order_id)
JOIN sql.store_products p USING(product_id)
GROUP BY 1
ORDER BY revenue DESC
```

Отношение количества клиентов по категориям:

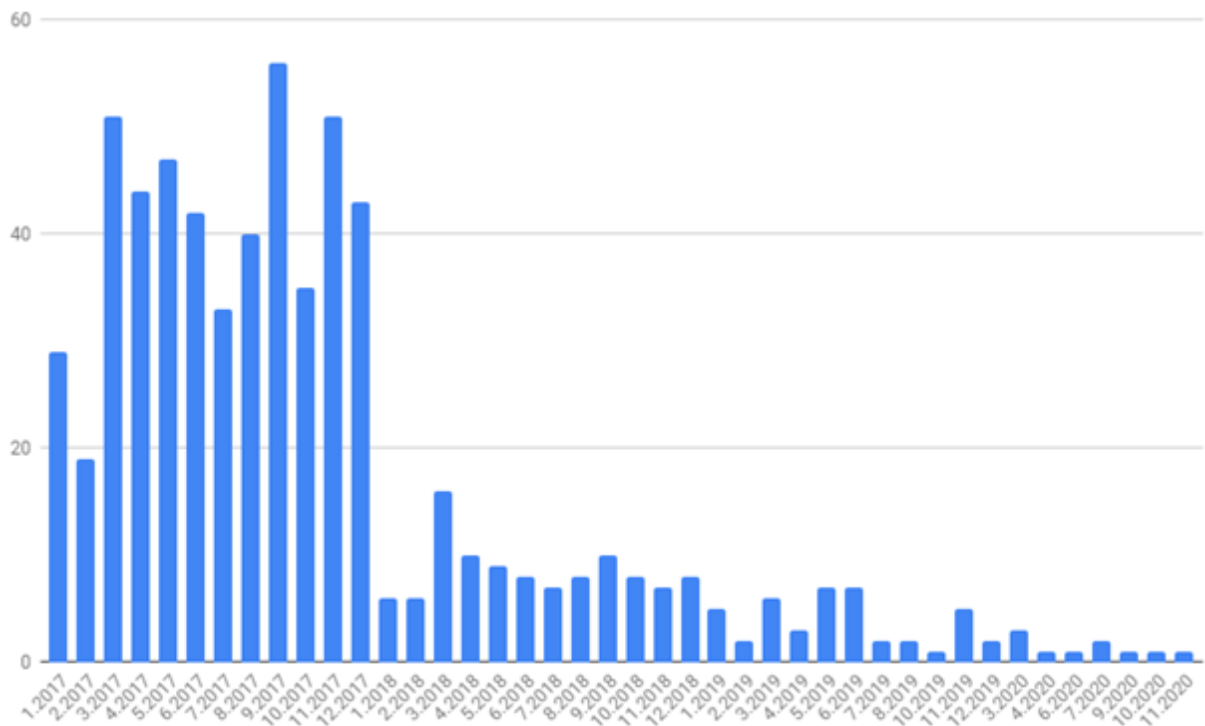


Отношение выручки по категориям:



2. Динамика новых B2B-клиентов по месяцам.

```
WITH min_order_date AS
(
  SELECT
    min(d.order_date) AS min_date,
    d.cust_id
  FROM sql.store_delivery d
  JOIN sql.store_customers c USING(cust_id)
  WHERE c.category = 'Corporate'
  GROUP BY d.cust_id)
SELECT
  date_trunc('month', m.min_date)::date date,
  count(m.cust_id) AS new_custs
FROM min_order_date m
GROUP BY 1
ORDER BY 1
```



В 2017 году было привлечено много клиентов, а после 2018 года новые клиенты почти не привлекались.

3. Основные показатели по корпоративным клиентам

Среднее количество различных товаров в заказах у корпоративных клиентов - 2,0

```
WITH order_count_product AS
(
  SELECT
    c.order_id,
    count(DISTINCT c.product_id) count_product
  FROM sql.store_carts c
  JOIN sql.store_delivery d USING(order_id)
  JOIN sql.store_customers cu USING(cust_id)
  WHERE cu.category='Corporate'
  GROUP BY 1)
SELECT
  ROUND(AVG(count_product), 1)
FROM order_count_product
```

Средняя сумма заказов у корпоративных клиентов - 285,9

```
WITH order_revenue AS
(
  SELECT d.order_id,
    SUM((p.price*c.quantity)*(1-c.discount)) revenue
  FROM sql.store_delivery d
  JOIN sql.store_carts c USING(order_id)
  JOIN sql.store_products p USING(product_id)
  JOIN sql.store_customers cu USING(cust_id)
  WHERE cu.category='Corporate'
  GROUP BY 1)
SELECT round(avg(revenue), 1)
FROM order_revenue
```

Среднее количество различных офисов у корпоративных клиентов - 6,2

```
WITH count_office AS
(
  SELECT DISTINCT d.cust_id,
    count(DISTINCT d.zip_code) zip_code
  FROM sql.store_delivery d
  JOIN sql.store_customers cu USING(cust_id)
  WHERE cu.category='Corporate'
  GROUP BY 1)
SELECT round(avg(zip_code), 1)
FROM count_office
```

4. Вывод

Основным источником дохода компании являются корпоративные клиенты. И по количеству, и по выручке B2B клиентов больше, чем B2C клиентов в 4,3 раза.

Необходимо заняться привлечением новых клиентов, в первую очередь B2C, так как средний чек B2C клиента (301,6) больше, чем средний чек B2B клиента (285,9)

А также необходимо поработать над средним количеством разных товаров в заказе. Сейчас это значение равно 2 - это маленький показатель.

Анализ логистики

1. Эффективность доставки

```
WITH ship_duration AS --рассчитывает длительность доставки каждого заказа
(
  SELECT
    order_id,
    ship_mode,
    order_date,
    ship_date,
    ship_date - order_date ship_duration
  FROM sql.store_delivery),
  plan_duration AS --добавляем столбец с планируемой
длительностью доставки для каждого типа доставки
(
  SELECT
    ship_mode,
    order_id,
    ship_duration,
    CASE
      WHEN ship_mode = 'Standard Class' THEN 6
      WHEN ship_mode = 'Second Class' THEN 4
      WHEN ship_mode = 'First Class' THEN 3
      ELSE 1
    END AS dil_class
  FROM ship_duration),
  late_orders AS --добавляем столбец с признаком «доставлен во
время» или «с опозданием» для каждого заказа
(
  SELECT
    ship_mode,
    order_id,
    ship_duration,
    dil_class,
    CASE
      WHEN ship_duration > dil_class THEN 'late'
      ELSE 'in_time'
    END AS late_orders
  FROM plan_duration),
  late_orders_cnt AS --рассчитывает для каждого типа доставки
общее количество заказов и количество заказов, доставленных не
вовремя
(
  SELECT
    ship_mode,
    count(order_id) orders_cnt,
    count(CASE
      WHEN late_orders = 'late' THEN order_id
```

```

        END) late_orders_cnt
    FROM late_orders
    GROUP BY 1)
SELECT
    ship_mode,
    orders_cnt,
    late_orders_cnt,
    round(((orders_cnt-late_orders_cnt)*100/orders_cnt),2) "% success"
FROM late_orders_cnt
ORDER BY 4

```

ship_mode	orders_cnt	late_orders_cnt	% success
Second Class	964	202	79.00
Standard Class	2994	309	89.00
First Class	787	1	99.00
Same Day	264	0	100.00

2. Доля заказов, отправленных вторым классом, которые были доставлены с опозданием, по кварталам.

```
WITH Second_Class_late AS
(
  SELECT
    order_id,
    order_date,
    CASE
      WHEN ship_date-order_date > 4 THEN 'late'
    END late_orders
  FROM sql.store_delivery
  WHERE ship_mode='Second Class')
SELECT
  date_trunc('quarter', order_date)::date QUARTER,
  count(late_orders)*100/count(order_id) "% late"
FROM Second_Class_late
GROUP BY 1
ORDER BY 1
```



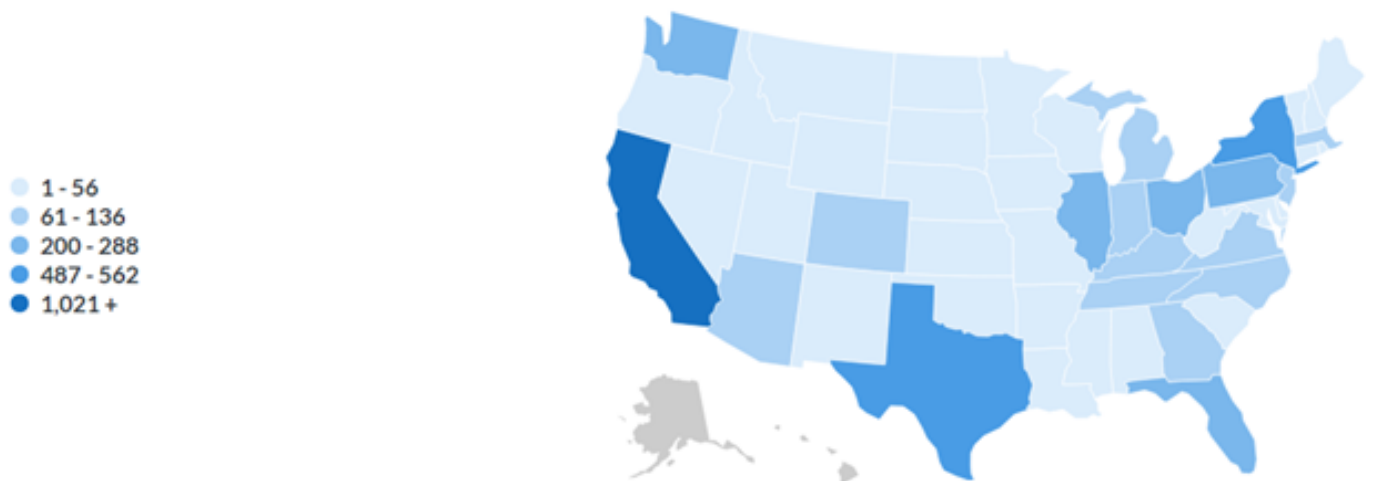
3. Распределение доставки по городам и штатам

наиболее популярный город по количеству доставок - New York City

```
SELECT
  city,
  count(DISTINCT order_id)
FROM sql.store_delivery
GROUP BY 1
ORDER BY 2 DESC
```

наиболее популярный штат по количеству доставок - California

```
SELECT
  state,
  count(DISTINCT order_id)
FROM sql.store_delivery
GROUP BY 1
ORDER BY 2 DESC
```



4. Вывод

Анализ логистики показал, что чаще всего с опозданием доходят заказы, отправленные вторым классом. При рассмотрении долей заказов, отправленных вторым классом, которые были доставлены с опозданием, по кварталам закономерностей не обнаружено, заказы доставляются с опозданием постоянно.

Наиболее популярный штат по количеству доставок - California. В этом штате и рекомендую открывать оффлайн-магазин, так как количество доставок в этот штат было почти вдвое больше, чем во второй по популярности штат New York.