**Анализ деятельности компании “Pens and Pencils”**

**ЦЕЛЬ:**

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

**ЗАДАЧИ:**

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).

Для проведения анализа деятельности компании представлены следующие данные:

* справочник клиентов;
* справочник товаров;
* таблица со списком товаров в заказе, их количества, а также скидкой для каждого товара;
* таблица с информацией о заказах и их доставке.

Заказы по предоставленным данным осуществлялись в период с 03.01.2017 г по 30.12.2020 г. За данный период было выполнено 5009 заказов на общую сумму

1 446 157 $ (с учетом скидок на определенный товар)

1. **Динамика продаж и распределение выручки по товарам**

**1.1** Проанализируем динамику распределения выручки компании по месяцам, для этого выведем необходимые данные (дату и объем выручки) и построим график, для лучшего визуального восприятия:

***Сумма выручки по месяцам:***

*select*

*date\_trunc ('month', sd.order\_date)::date as date,*

*round (sum ((sp.price - sp.price \* sc.discount) \* sc.quantity)) as revenue*

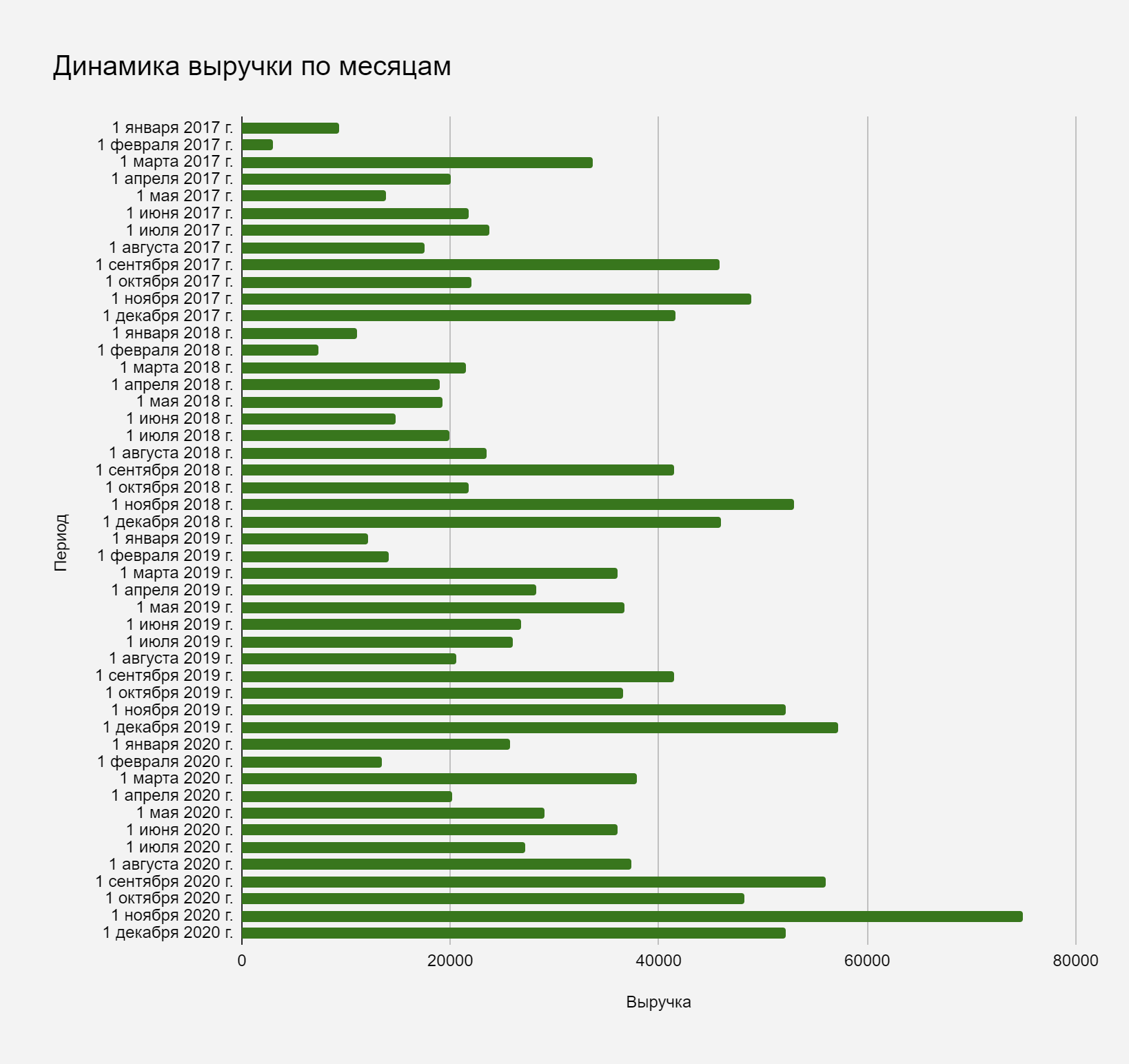
*from sql.store\_carts as sc*

*join sql.store\_products as sp on sc.product\_id = sp.product\_id*

*join sql.store\_delivery as sd on sd.order\_id = sc.order\_id*

*group by date*

*order by date*



Исходя из полученных данных, можно заметить, что доход компании в целом растет, но наблюдается рост и падение выручки в зависимости от сезона.

**1.2** Определим какие товары пользуются наибольшим спросом, для этого составим рейтинг продаж различных товаров по различным категориям и подкатегориям:

***Сумма выручки по различным категориям и подкатегориям:***

*select*

*sp.category,*

*sp.subcategory,*

*round(sum ((sp.price - sp.price \* sc.discount) \* sc.quantity)) as revenue*

*from sql.store\_carts as sc*

*join sql.store\_products as sp on sc.product\_id = sp.product\_id*

*group by category, subcategory*

*order by revenue desc*

Результат запроса:

|  |  |  |
| --- | --- | --- |
| **категория товара** | **подкатегория товара** | **объем выручки** |
| Furniture | Chairs | 235318 |
| Technology | Phones | 221110 |
| Office Supplies | Storage | 179736 |
| Technology | Accessories | 125442 |
| Furniture | Tables | 114532 |
| Technology | Machines | 113978 |
| Technology | Copiers | 95844 |
| Office Supplies | Paper | 62148 |
| Furniture | Bookcases | 59271 |
| Office Supplies | Appliances | 55550 |
| Furniture | Furnishings | 51695 |
| Office Supplies | Binders | 45312 |
| Office Supplies | Supplies | 37457 |
| Office Supplies | Art | 22148 |
| Office Supplies | Envelopes | 13809 |
| Office Supplies | Labels | 10345 |
| Office Supplies | Fasteners | 2459 |

Из таблицы видим, что в тройку лидеров продаж входят кресла, телефоны и места для хранения офисных принадлежностей.

**1.3** Рассмотрим, какие из товаров приносят больше всего прибыли и какую долю от общей выручки составляют их продажи. Для этого составим рейтинг топ-25 товаров по объему выручки, а также посчитаем количество проданных товаров и их долю от общей выручки в процентах:

***Топ-25 товаров по объему выручки:***

*select*

*sp.product\_nm,*

*round ((sum (quantity \* price \* (1 - discount))), 2) as revenue,*

*sum (sc.quantity) as quantity,*

*round ((sum (quantity \* price \* (1 - discount)) / (select*

*sum (quantity \* price \* (1 - discount))*

*from sql.store\_products as p*

*join sql.store\_carts as c on p.product\_id = c.product\_id) \* 100), 2) as percent\_from\_total*

*from sql.store\_carts as sc*

*join sql.store\_products as sp on sc.product\_id = sp.product\_id*

*group by product\_nm*

*order by revenue desc*

*limit 25*

Результат запроса:

|  |  |  |  |
| --- | --- | --- | --- |
| **наименование товара** | **объем выручки** | **количество проданных товаров** | **доля от общей выручки, %** |
| Canon imageCLASS 2200 Advanced Copier | 36959,82 | 20 | 2,56 |
| HON 5400 Series Task Chairs for Big and Tall | 15309,53 | 39 | 1,06 |
| 3D Systems Cube Printer, 2nd Generation, Magenta | 14299,89 | 11 | 0,99 |
| High Speed Automatic Electric Letter Opener | 13624,21 | 11 | 0,94 |
| Martin Yale Chadless Opener Electric Letter Opener | 13325 | 22 | 0,92 |
| Cisco TelePresence System EX90 Videoconferencing Unit | 11319,24 | 6 | 0,78 |
| Hewlett Packard LaserJet 3310 Copier | 11303,69 | 38 | 0,78 |
| Samsung Galaxy Mega 6.3 | 11154,87 | 39 | 0,77 |
| Canon PC1060 Personal Laser Copier | 9295,83 | 19 | 0,64 |
| HP Designjet T520 Inkjet Large Format Printer - 24" Color | 9187,5 | 12 | 0,64 |
| Global Troy Executive Leather Low-Back Tilter | 9082,87 | 32 | 0,63 |
| Honeywell Enviracaire Portable HEPA Air Cleaner for 17' x 22' Room | 9043,55 | 40 | 0,63 |
| Tennsco 6- and 18-Compartment Lockers | 8740,17 | 43 | 0,6 |
| Plantronics CS510 - Over-the-Head monaural Wireless Headset System | 8657,89 | 35 | 0,6 |
| SAFCO Arco Folding Chair | 8100,95 | 53 | 0,56 |
| Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish | 7805,48 | 24 | 0,54 |
| Apple iPhone 5 | 7798 | 24 | 0,54 |
| Tennsco Double-Tier Lockers | 7704,86 | 47 | 0,53 |
| Ativa V4110MDD Micro-Cut Shredder | 7699,89 | 11 | 0,53 |
| Hon Deluxe Fabric Upholstered Stacking Chairs, Rounded Back | 7446,44 | 51 | 0,51 |
| Tennsco Single-Tier Lockers | 7326,59 | 28 | 0,51 |
| Office Star - Professional Matrix Back Chair with 2-to-1 Synchro Tilt and Mesh Fabric Seat | 7019,5 | 28 | 0,49 |
| Zebra ZM400 Thermal Label Printer | 6965,7 | 6 | 0,48 |
| Hot File 7-Pocket, Floor Stand | 6510,77 | 49 | 0,45 |
| Bretford Rectangular Conference Table Tops | 6497,82 | 46 | 0,45 |

1. **Портрет клиента**

**2.1** Определим, сколько у компании *B2B (Corporate)* и *B2C (Consumer)* клиентов и какую долю от общей выручки они приносят.

Для этого вычислим количество клиентов и выручку по категориям клиента.

***Количество клиентов и выручка по категориям клиента:***

*select*

*scu.category,*

*count (distinct scu.cust\_id) as cust\_cnt,*

*round(sum((sp.price - sp.price \* sc.discount) \* sc.quantity)) as revenue*

*from sql.store\_delivery as sd*

*join sql.store\_customers as scu on sd.cust\_id = scu.cust\_id*

*join sql.store\_carts as sc on sd.order\_id = sc.order\_id*

*join sql.store\_products as sp on sp.product\_id = sc.product\_id*

*group by scu.category*

*order by revenue desc*

Результат запроса:

|  |  |  |
| --- | --- | --- |
| **категория** | **количество клиентов** | **прибыль** |
| Corporate | 645 | 1172009 |
| Consumer | 148 | 274148 |

По полученным данным видно, что корпоративных клиентов больше и выручка от них выше, чем от розничных покупателей.

**2.2** Проанализируем динамику новых *B2B*-клиентов по месяцам. Заодно выясним, выручка растёт за счёт увеличения продаж по старым клиентам или же за счёт привлечения новых.

***Количество новых корпоративных клиентов по месяцам:***

*select*

*date\_trunc ('month', date)::date as month,*

*count (cust\_id) as new\_custs*

*from (*

*select*

*min (sd.order\_date) as date,*

*scu.cust\_id*

*from sql.store\_delivery as sd*

*join sql.store\_customers as scu on sd.cust\_id = scu.cust\_id*

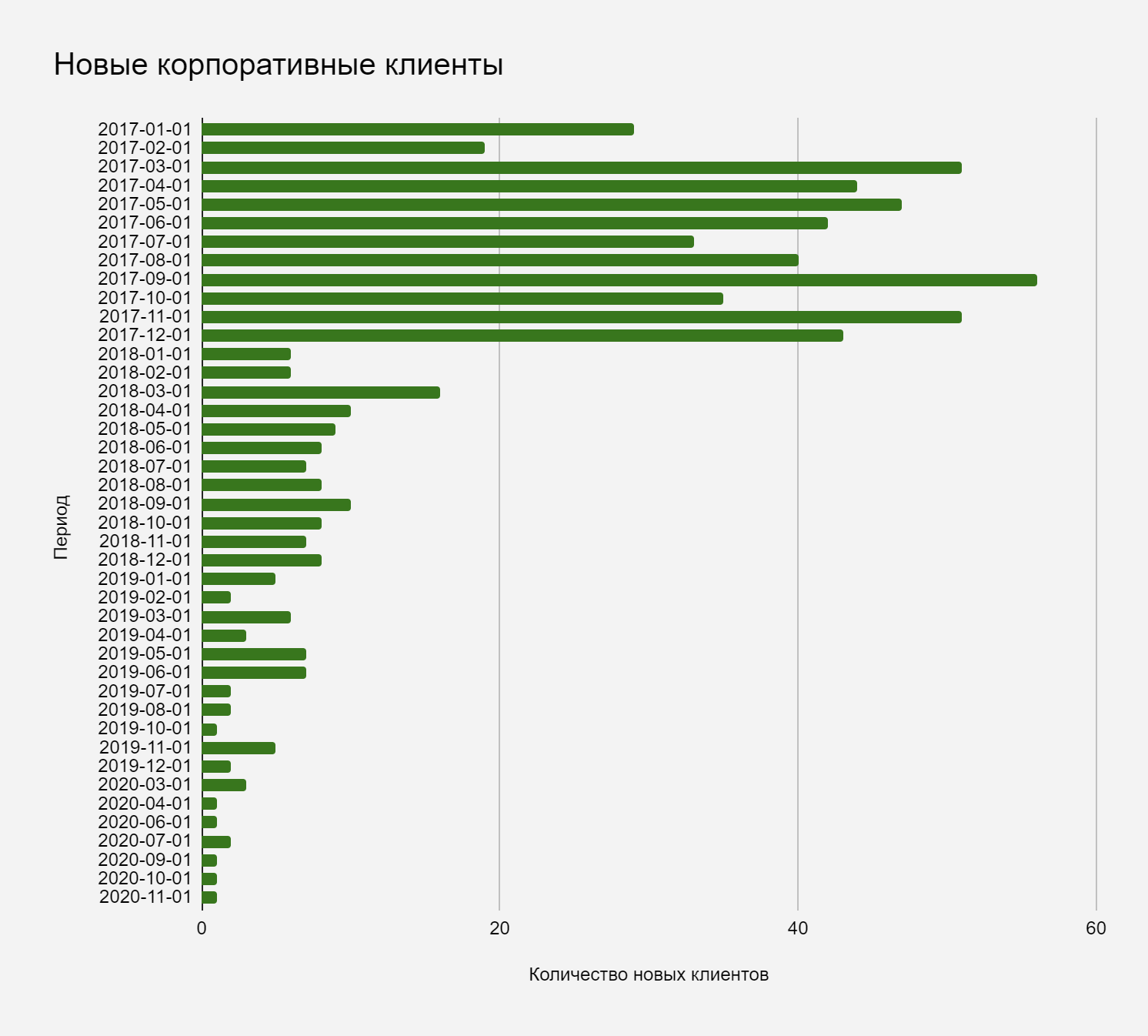
*where scu.category = 'Corporate'*

*group by scu.cust\_id) as cust\_date*

*group by month*

*order by month*

Результат запроса:

******

Из диаграммы, видно, что количество новых клиентов уменьшается, а значит выручка компании растет за счет старых клиентов.

**2.3** Изучим основные показатели по корпоративным клиентам, для этого вычислим каково среднее количество различных товаров в заказах, среднюю сумму заказов, среднее количество различных офисов у корпоративных клиентов:

***-- среднее количество различных товаров в заказах и средняя сумма заказов у корпоративных клиентов:***

*with quantity\_and\_sum\_order as (*

*select*

*distinct scu.cust\_nm as cust\_nm,*

*count (distinct sc.order\_id) as order\_id,*

*count (distinct sp.product\_nm) as qnt,*

*sum ((sp.price - sp.price \* sc.discount) \* sc.quantity) as revenue*

*from sql.store\_delivery as sd*

*join sql.store\_customers as scu on sd.cust\_id = scu.cust\_id*

*join sql.store\_carts as sc on sd.order\_id = sc.order\_id*

*join sql.store\_products as sp on sp.product\_id = sc.product\_id*

*where scu.category = 'Corporate'*

*group by cust\_nm*

*order by cust\_nm),*

***-- среднее количество различных офисов у корпоративных клиентов:***

*qnt\_office as*

*(select*

*distinct scu.cust\_nm as cust\_nm,*

*count (distinct sd.zip\_code) as cnt\_office*

*from sql.store\_delivery as sd*

*join sql.store\_customers as scu on sd.cust\_id = scu.cust\_id*

*where scu.category = 'Corporate'*

*group by cust\_nm*

*order by cust\_nm)*

*select*

*round (sum (qnt) / sum (order\_id), 1) as avg\_qnt\_product,*

*round (sum (revenue) / sum (order\_id), 1) as avg\_sum\_order,*

*round (avg (cnt\_office), 1) as avg\_cnt\_office*

*from quantity\_and\_sum\_order as qs*

*join qnt\_office as qo on qs.cust\_nm = qo.cust\_nm*

Результат запроса:

|  |  |  |
| --- | --- | --- |
| **среднее количество товаров в заказе** | **средняя сумма заказов** | **среднее количество различных офисов** |
| 2.0 | 285.9 | 6.2 |

Исходя из полученных данных, можно сделать вывод, что наибольшую ценность для компании представляют корпоративные клиенты, потому что на них приходится большая доля выручки, но число новых клиентов неуклонно падает, а так как количество и сумма товаров в заказе в среднем весьма невысокая, то, соответственно, и выручка будет падать. В таком случае необходимо предпринять действия по привлечению новых корпоративных клиентов и увеличению их среднего чека.

1. **Логистика компании**

**3.1** Для того чтобы оценить текущую картину по логистике доставок, необходимо определить насколько эффективно выполняются заказы и как распределяются доставки и выручка по штатам и городам. Для этого определим, какая доля заказов выполняется в срок.

***Тип доставки, общее количество заказов, количество заказов которые не были доставлены вовремя и доля выполненных вовремя заказов:***

*with delivery\_time as*

*(select*

*ship\_mode,*

*count (order\_id)::numeric as orders\_cnt,*

*count (case when ship\_mode = 'Standard Class' and (ship\_date - order\_date) > 6 then order\_id*

*when ship\_mode = 'Second Class' and (ship\_date - order\_date) > 4 then order\_id*

*when ship\_mode = 'First Class' and (ship\_date - order\_date) > 3 then order\_id*

*when ship\_mode = 'Same Day' and (ship\_date - order\_date) > 0 then order\_id*

*end)::numeric as late\_orders\_cnt*

*from sql.store\_delivery*

*group by ship\_mode)*

*select*

*ship\_mode,*

*orders\_cnt,*

*late\_orders\_cnt,*

*ROUND ((orders\_cnt - late\_orders\_cnt)\*100 / orders\_cnt, 2) AS "% success"*

*from delivery\_time*

*order by "% success"*

Результат запроса

|  |  |  |  |
| --- | --- | --- | --- |
| **тип доставки** | **общее количество заказов** | **количество заказов, которые не были доставлены вовремя** | **доля выполненных вовремя заказов, %** |
| Second Class | 964 | 202 | 79,05 |
| Standard Class | 2994 | 309 | 89,68 |
| Same Day | 264 | 12 | 95,45 |
| First Class | 787 | 1 | 99,87 |

По результатам полученных данных можно сделать вывод, что чаще всего с опозданием доходят заказы, отправленные вторым классом (*Second Class*)

**3.2** Необходимо выявить насколько систематично это происходит. Возможно, у службы доставки были проблемы только на протяжении ограниченного периода. Для этого вычислим ***долю заказов, отправленных вторым классом, которые были доставлены с опозданием, по кварталам:***

*with delivery\_time\_sc as*

*(select*

*ship\_mode,*

*date\_trunc ('quarter', order\_date)::date as quarter,*

*count (order\_id)::numeric as orders\_cnt,*

*count (case when (ship\_date - order\_date) > 4 then order\_id*

*end)::numeric as late\_orders\_cnt*

*from sql.store\_delivery*

*group by ship\_mode, quarter)*

*select*

*quarter,*

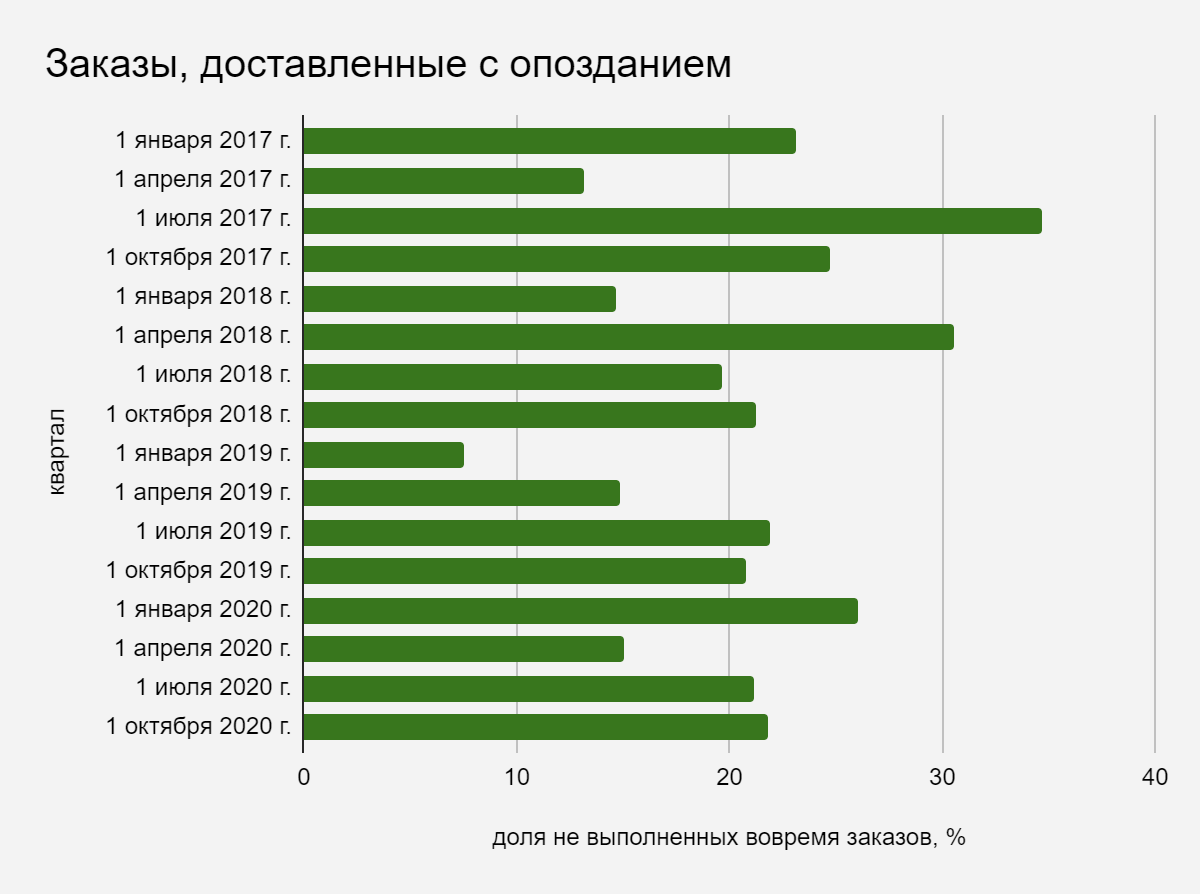
*ROUND ((late\_orders\_cnt)\*100 / orders\_cnt, 2) AS "% unsuccess"*

*from delivery\_time\_sc*

*where ship\_mode = 'Second Class'*

*order by quarter*

Результат запроса:



Из графика видно, что заказы отправленные “Second Class” систематически доставляются с опозданием.

**3.3** Проанализируем, в какой штат и город осуществляется больше всего доставок.

***Вычислим количество доставок по штатам и визуализируйте результат с помощью карты:***

*select*

*state,*

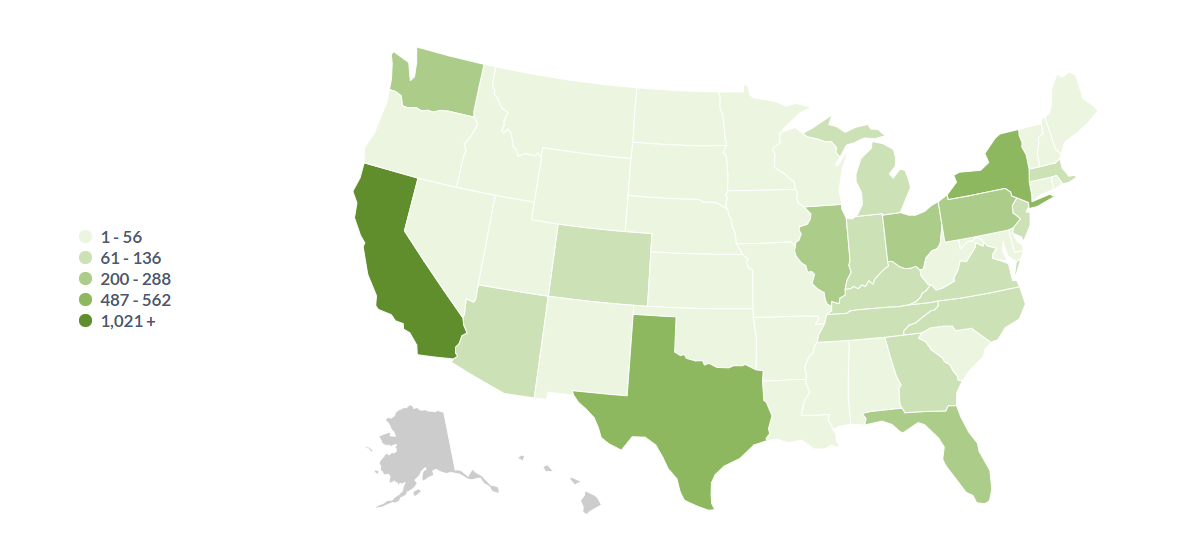
*count(distinct order\_id)*

*from sql.store\_delivery*

*group by 1*

*order by 2 desc*

Результат запроса:



Самый популярный город по количеству доставок является **New York City**,450 заказов было доставлено, но на весь штат **New York** приходится только 562 доставки, в то время как **Los Angeles** и **San Francisco** занимают 2е и 3е места в рейтинге с количеством 384 и 265 доставок, соответственно, а по всему штату **California** было осуществлено 1021 доставка, что является более ⅕ всех доставок по стране.

**Вывод**

На основе проведенного анализа деятельности компании “*Pens and Pencils”*, можно дать следующие рекомендации:

* обратить внимание на убывающую динамику новых клиентов и принять меры для ее роста;
* разработать стратегию по увеличению среднего чека покупателя;
* проанализировать логистику компании, выяснить, почему происходят задержки в доставках;

На данный момент у компании есть только склад, откуда отправляются все товары, — находится он в городе Хьюстон, штат Техас (*Houston*, *Texas*) и, чтобы разгрузить логистику, снизить стоимость доставки и привлечь новых клиентов, рекомендую масштабировать бизнес, а именно, открыть офлайн-магазин в штате Калифорния.