

Студент: Минин Владислав

Группа: DAPR-210

Почта: vminin2017@mail.ru

**Задание 27.8.1 (PJ-04) PROJECT-2. Решение бизнес-задач с помощью SQL**

**Задача:**

Проанализировать работу компании Pens and Pencils с точки зрения ее эффективности и дать рекомендации по масштабированию бизнеса, предложить штат, в котором лучше открыть офлайн-магазин

**Формализация задачи:**

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)

## 

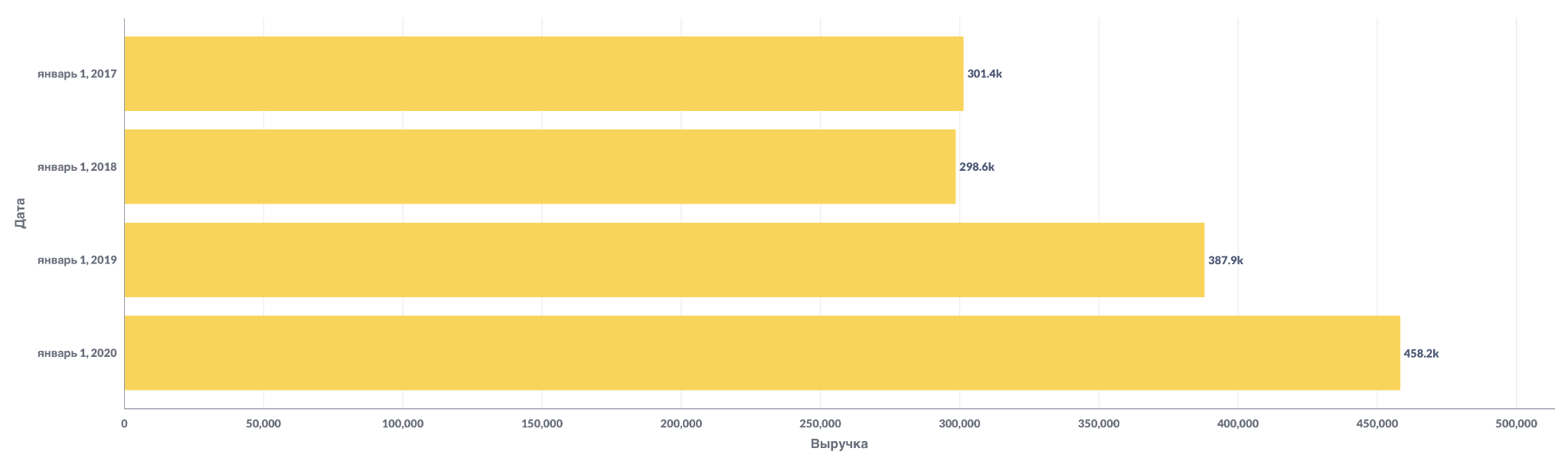
## 

## 

# **1) Эффективность продаж**

**Отслеживание динамики продаж и выручки:**

Для того, чтобы более детально говорить о прибыльности компании и о ее положении дел, рассмотрим выручку с начала 2017 года: на графике четко прослеживается тенденция роста выручки на всем временном промежутке. Общая картина говорит о том, что суммарно каждый год компания усиливает свое финансовое положение.



*SELECT*

*DATE\_TRUNC('year', order\_date)::date as "date",*

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

*FROM sql.store\_products p*

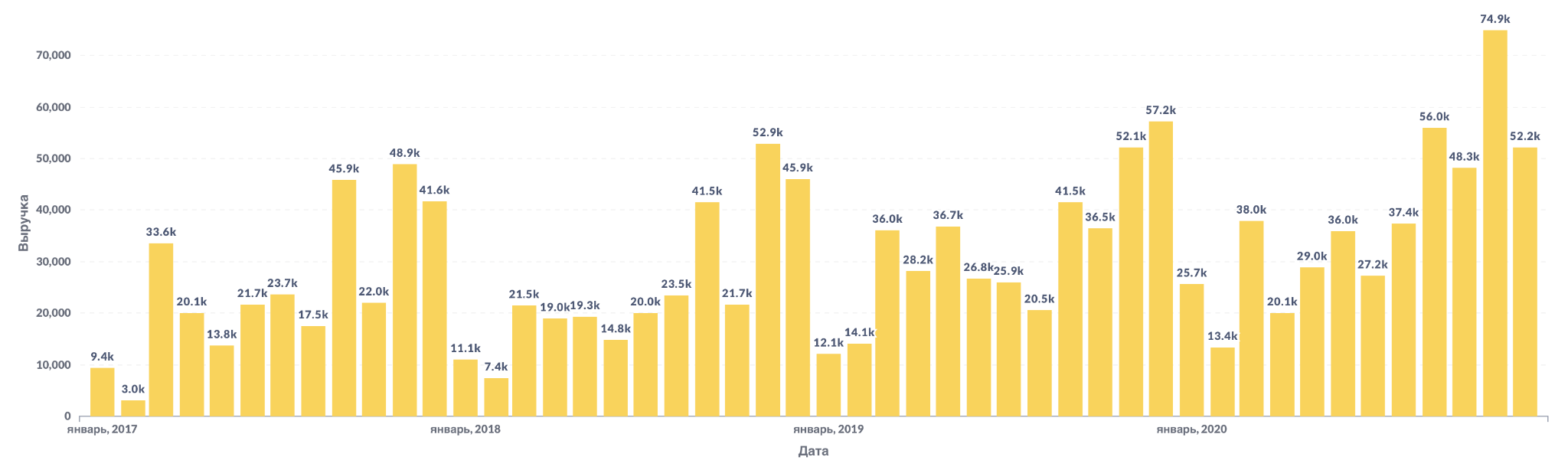
*JOIN sql.store\_carts c ON p.product\_id = c.product\_id*

*JOIN sql.store\_delivery d ON d.order\_id = c.order\_id*

*GROUP BY "date"*

*ORDER BY "date"*

Более подробно распределение выручки можно увидеть на графике ниже помесячно. Исходя из него можно сделать более детальный вывод о том, что в движении денежных средств присутствует сезонность - в январе каждого года происходит падение выручки со стабильным последующим ростом, начиная со следующих пары месяцев.



*SELECT*

*DATE\_TRUNC('month', order\_date)::date as "date",*

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

*FROM sql.store\_products p*

*JOIN sql.store\_carts c ON p.product\_id = c.product\_id*

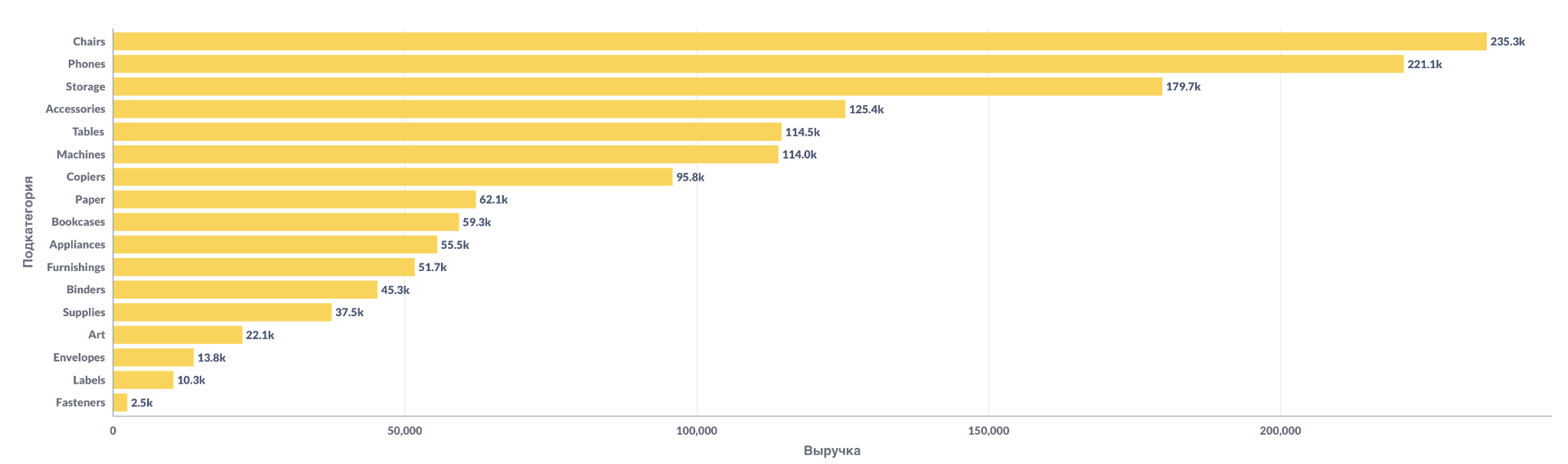
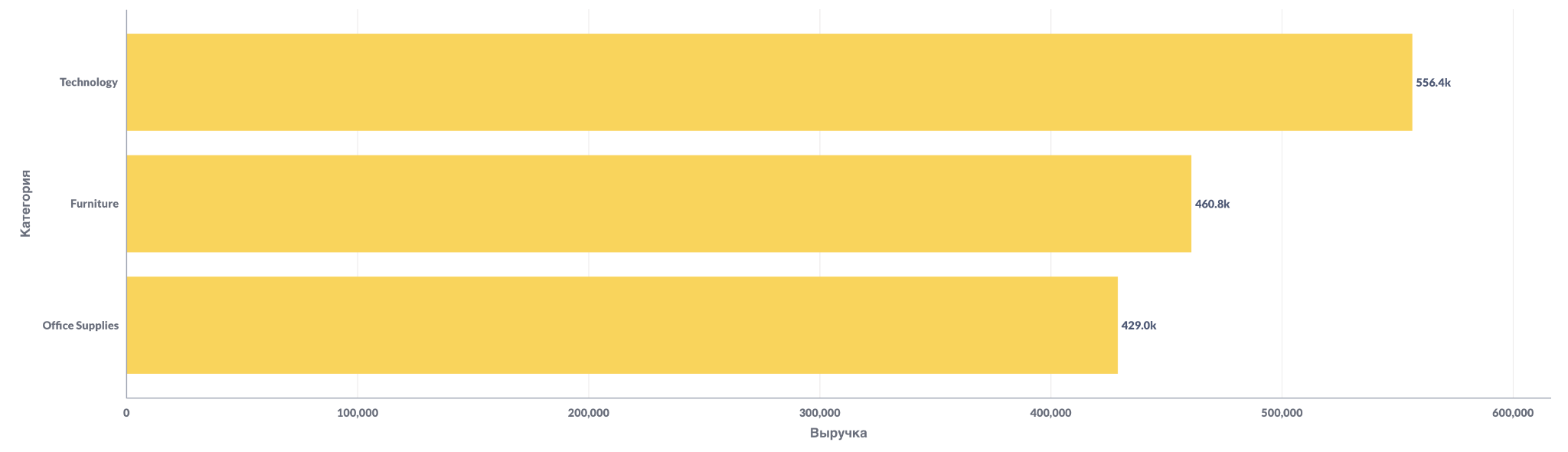
*JOIN sql.store\_delivery d ON d.order\_id = c.order\_id*

*GROUP BY "date"*

*ORDER BY "date"*

**Выручка по категориям и подкатегориям товаров:**

Теперь стоит рассмотреть, продажи каких офисных товаров увенчались большим успехом и составляют основную долю выручки. Товары в компании подразделены на 3 категории и 17 подкатегорий. Среди трех основных категорий 556 000$ приходятся на продажу оргтехники (Technology), что обуславливает ее основную долю в общей выручке, затем идет категория мебели (Furniture) с выручкой в 460 000$ и канцтовары (Office Supplies) - 429 000$. Более подробное распределение выручки с продаж можно посмотреть на втором графике. Топ-3 подкатегорий товаров приходят из разных категорий. Chairs из категории Furniture являются основным товаром - на них приходится 235 000$, Phones занимают основное место в категории Technology - на них приходится 221 000$, и замыкает троицу Storage из категории Office Supplies - с выручкой в 179 000$.



*SELECT*

*p.category category,*

*p.subcategory subcategory,*

*ROUND(SUM(p.price\*c.quantity\*(1 - c.discount))) revenue*

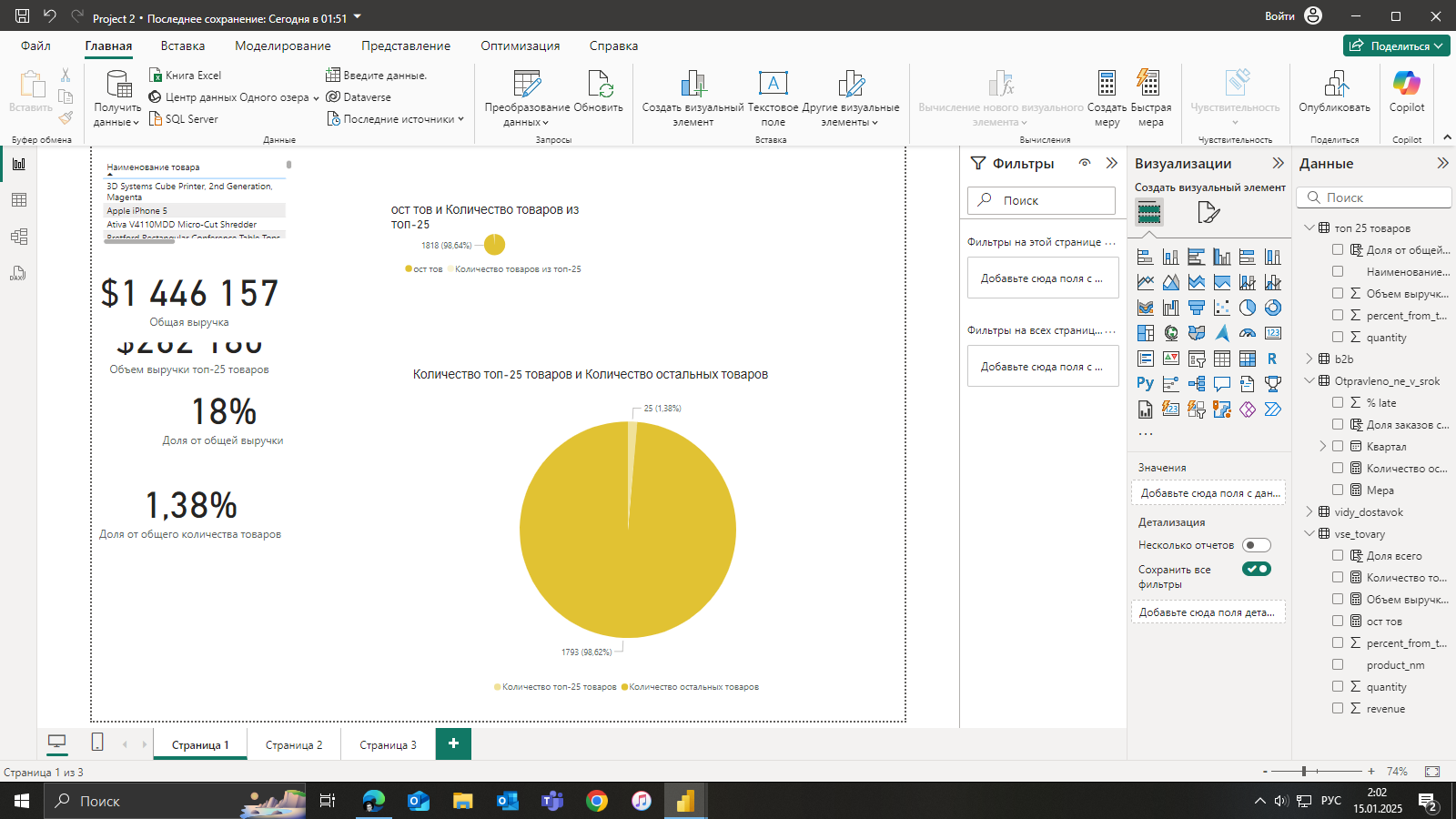
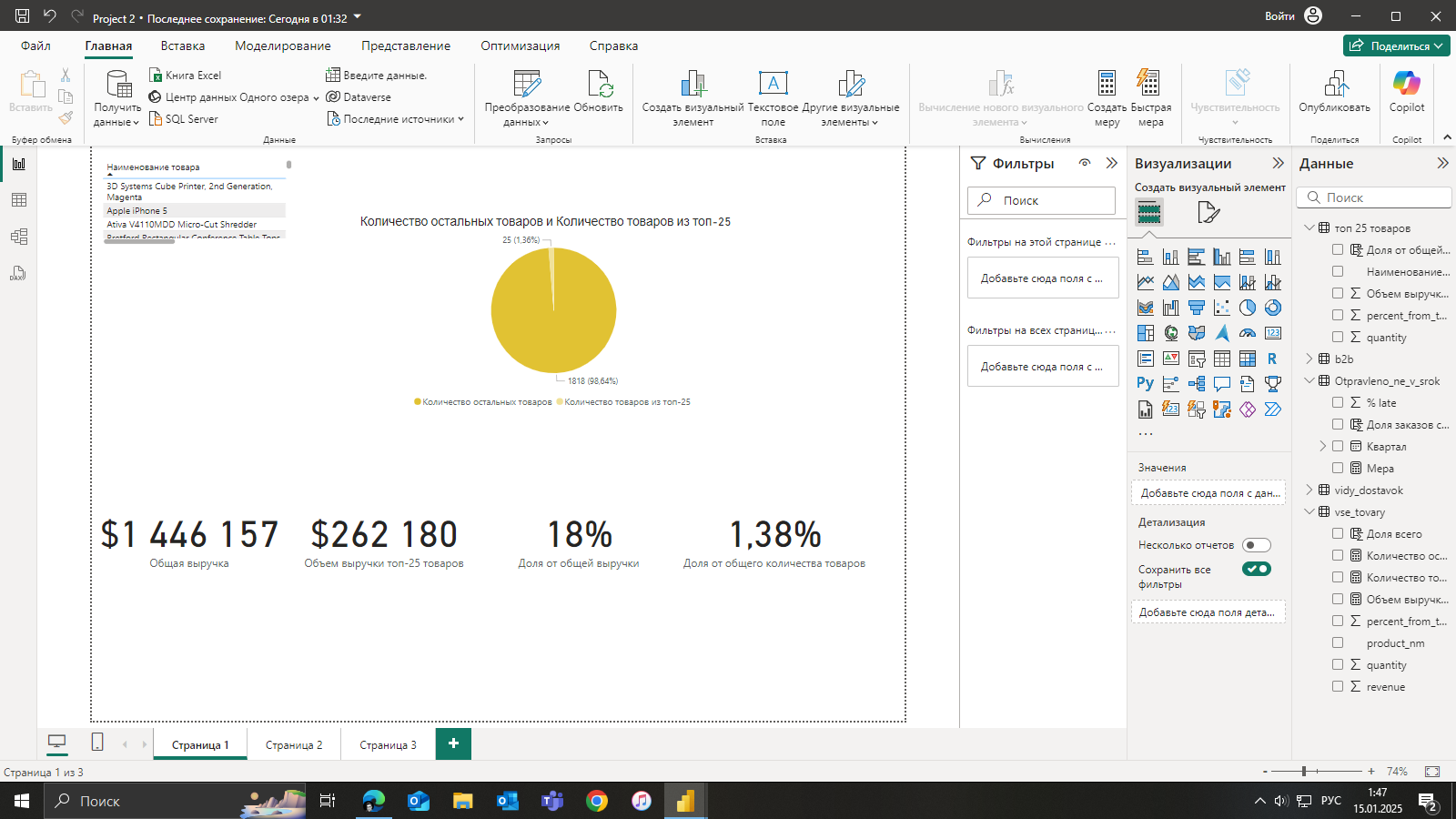
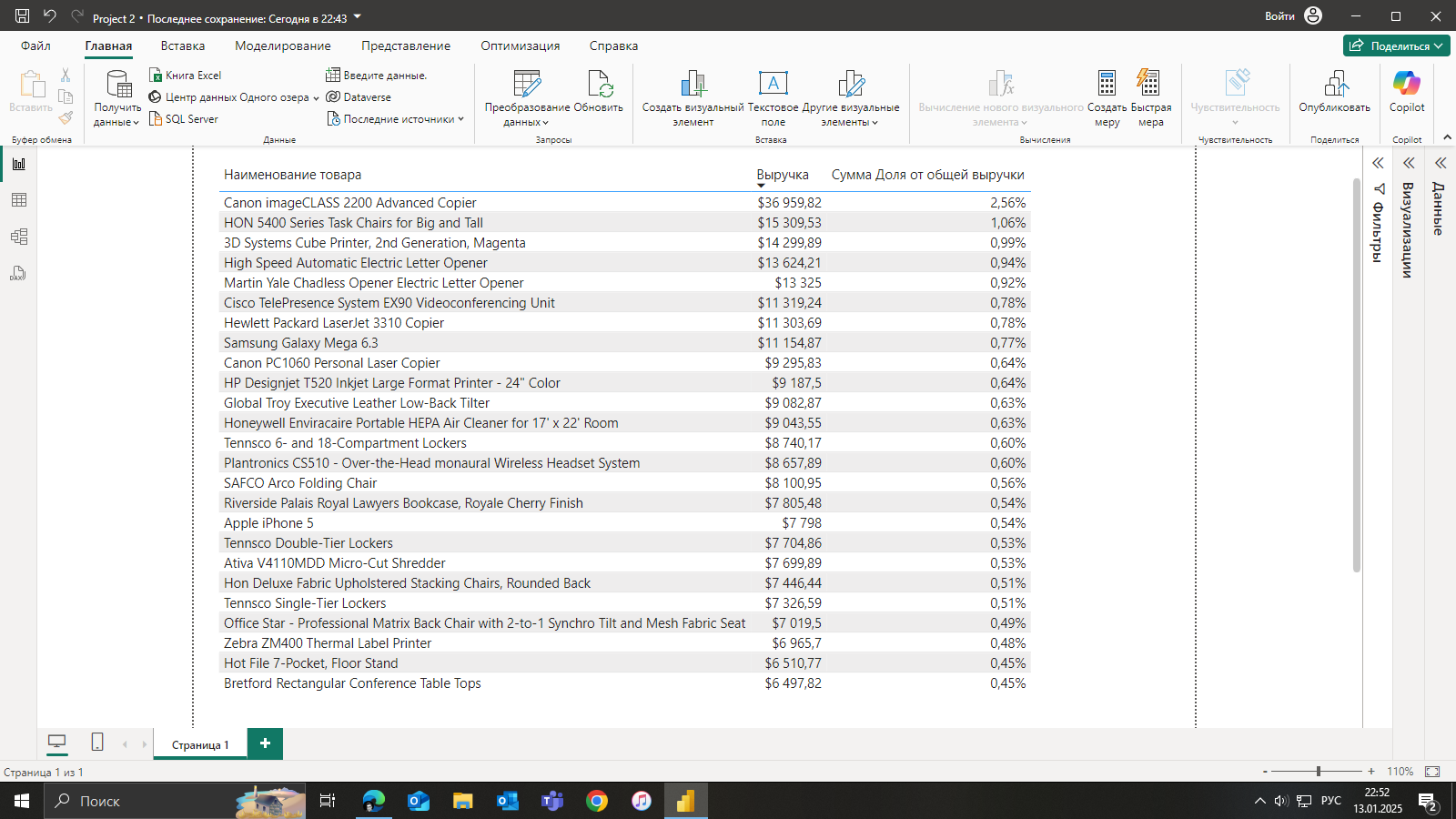
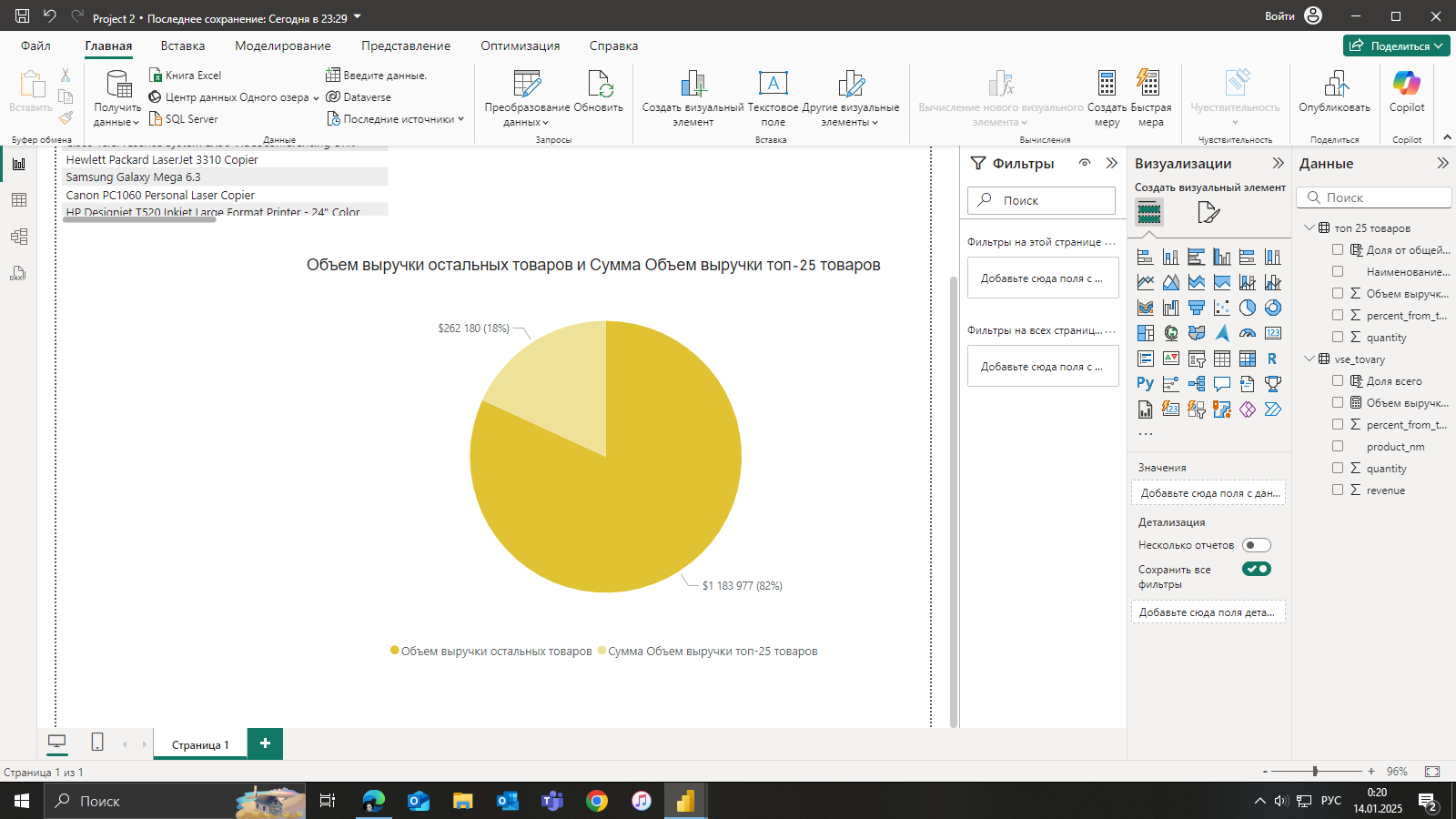
*FROM sql.store\_products p*

*JOIN sql.store\_carts c ON p.product\_id = c.product\_id*

*GROUP BY category, subcategory*

*ORDER BY revenue DESC*

**Самые продаваемые товары и их доля от общей выручки:**

Компания продает 1818 различных товаров. Для того, чтобы узнать, какие товарные наименования занимают лидирующие позиции, был выделен топ-25 самых продаваемых товаров. В таблице приведен перечень этих товаров с выручкой, которая на них приходится, по убыванию, и с долей от общей выручки, которую забирают на себя эти товары. Можно заметить, что этот топ-25 товаров (что составляет немногим более 1% от общего количества товаров) генерирует 18% общей выручки компании (более 260 000$), что свидетельствует об очень большом спросе на них со стороны партнеров и потребителей.

*WITH total\_revenue AS*

*(SELECT*

*ROUND(SUM(c.quantity \* p.price \* (1 - c.discount))) total*

*FROM sql.store\_carts c*

*JOIN sql.store\_products p on c.product\_id = p.product\_id)*

*SELECT*

*product\_nm,*

*ROUND(SUM(c.quantity \* p.price \* (1 - c.discount)), 2) revenue,*

*SUM(quantity) quantity,*

*ROUND(SUM(c.quantity \* p.price \* (1 - c.discount)) / total \* 100, 2) percent\_from\_total*

*FROM sql.store\_carts c*

*JOIN sql.store\_products p ON c.product\_id = p.product\_id*

*JOIN total\_revenue ON TRUE*

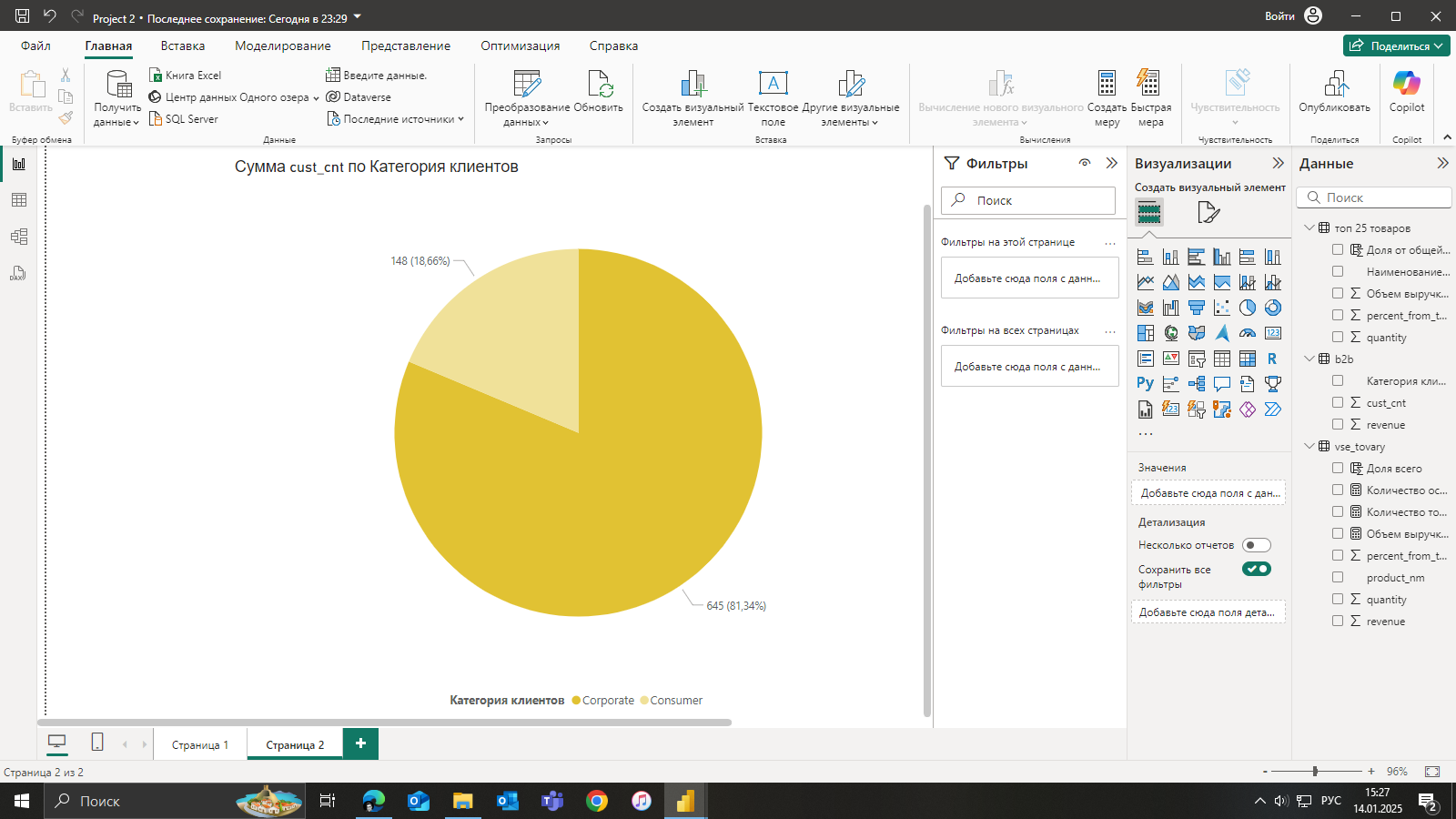
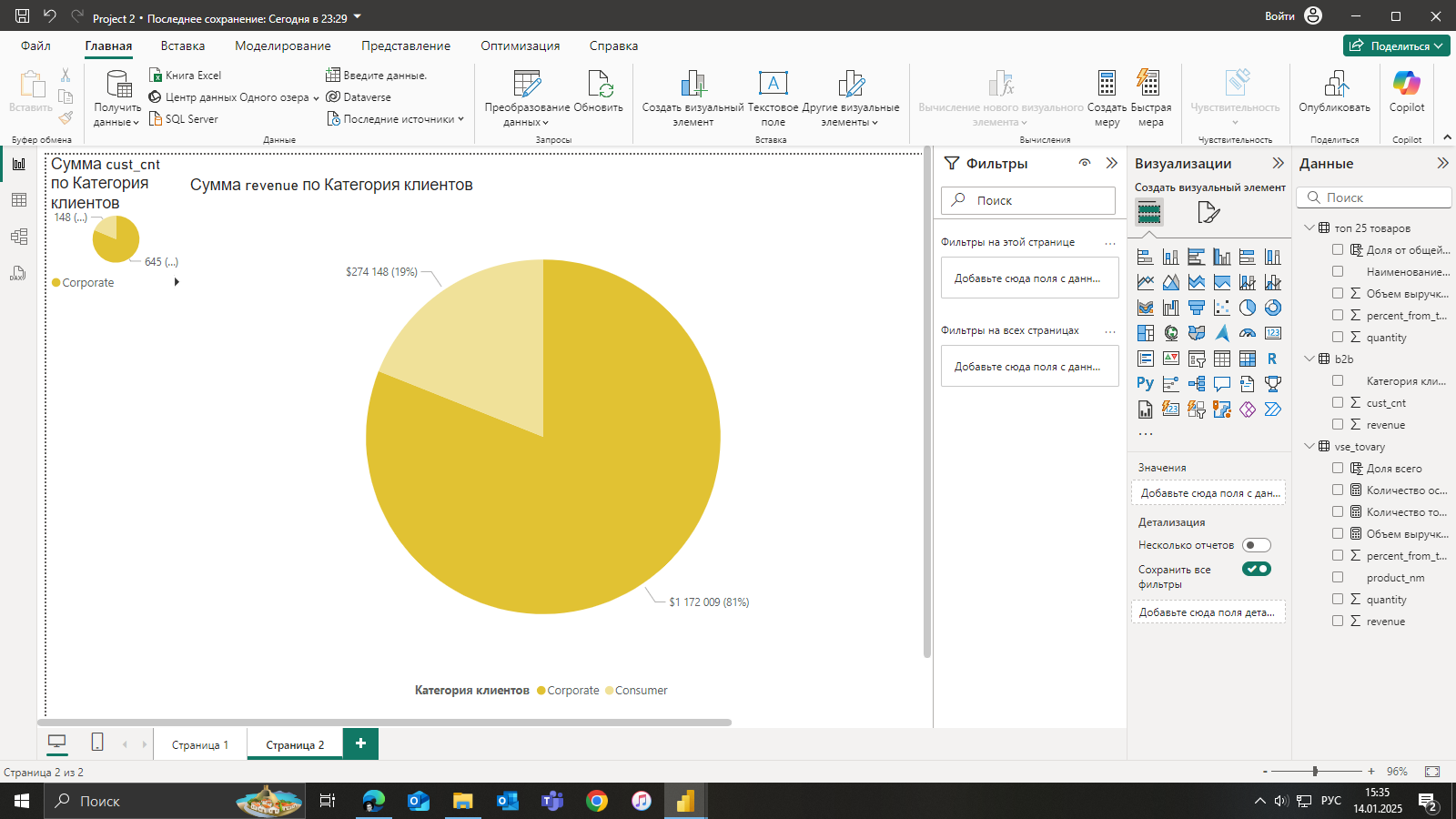
*GROUP BY product\_nm, total*

*ORDER BY percent\_from\_total DESC*

*LIMIT 25*

# **2) Исследование клиентов компании**

**Выручка клиентов по категориям:**

Pens and Pencils является одновременно и B2B, и B2C компанией, поэтому важно разделить эти две категории и посмотреть, как между ними распределяется выручка в денежном и процентном соотношении. Доля B2C клиентов (Consumer) составляет 18,66% или 148 человек, в то время как на долю B2B клиентов (Corporate) приходится 81,34% - 645 человек. Примерно такое же соотношение наблюдается и в распределении выручки: B2C клиенты приносят 19% всей выручки (274 148$), когда B2B клиенты приносят 81% (1 172 009$). 

Исходя из этого можно сделать вывод, что основным направлением деятельности компании является B2B сектор, и именно на него следует больше опираться при дальнейшем анализе, принимая во внимание все особенности характеристики этого сегмента.

*SELECT*

*cu.category,*

*COUNT(DISTINCT cu.cust\_id) cust\_cnt,*

*ROUND(SUM((p.price \* ca.quantity) \* (1 - ca.discount))) revenue*

*FROM sql.store\_customers cu*

*JOIN sql.store\_delivery d ON cu.cust\_id = d.cust\_id*

*JOIN sql.store\_carts ca ON d.order\_id = ca.order\_id*

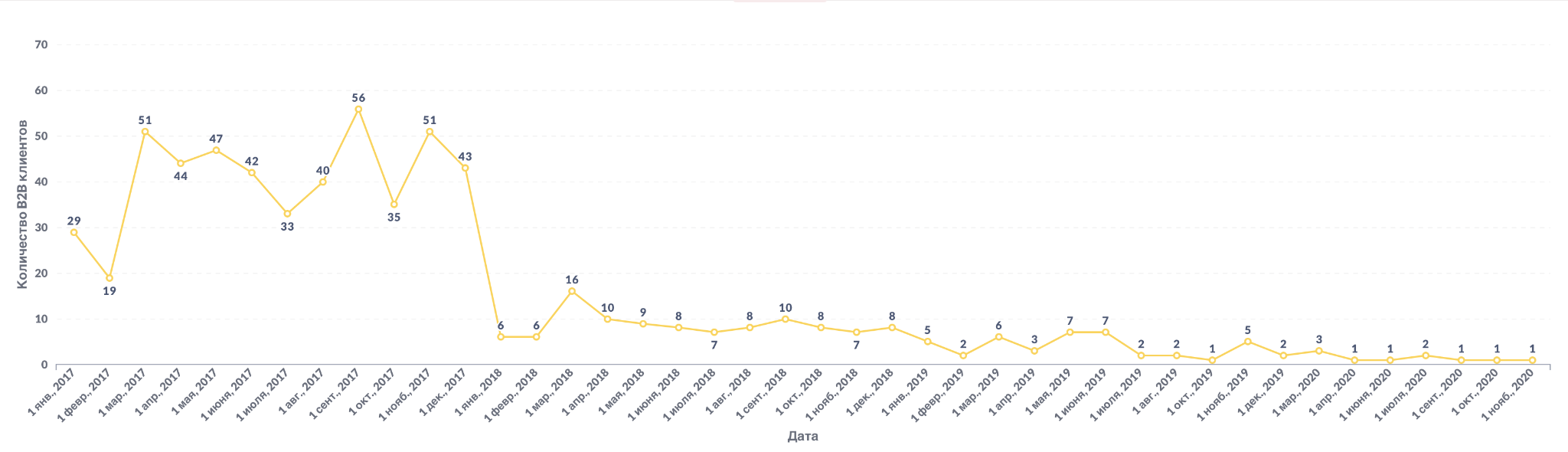
*JOIN sql.store\_products p ON ca.product\_id = p.product\_id*

*GROUP BY cu.category*

*ORDER BY revenue DESC*

**Особенности удержания и привлечения B2B клиентов:**

Сделав вывод о том, что основными клиентами компании являются клиенты из B2B сектора, можно поподробнее поговорить об их характеристиках, и их поведении на всем протяжении временного промежутка. Таким образом, можно посмотреть сколько клиентов у компании было в разные временные точки, был ли рост в количестве таких клиентов или спад, то есть насколько хорошо развито удержание старых и привлечение новых клиентов. На графике можно четко увидеть флуктуации в количестве привлеченных корпоративных клиентов на протяжении всего 2017 года, но можно с уверенностью сказать, что их количество было выше 30 в разные временные точки и оставалось довольно высоким. Однако, резкий спад случился в январе 2018 - в самом начале года. 1 декабря 2017 было привлечено 43 клиента, когда в январе следующего года их стало всего лишь 6. Более того, такая тенденция прослеживается на всем остальном таймлайне. Следует отметить, что несмотря на это выручка компании увеличивается из года в год, а значит можно сделать вывод, что ее приносят уже привлеченные клиенты в 2017 году, но следует исключить возможность того, что это происходит за счет B2C клиентов.



*WITH first\_order\_month AS*

*(SELECT*

*MIN(DATE\_TRUNC('month', order\_date))::date "month",*

*d.cust\_id*

*FROM sql.store\_delivery d*

*JOIN sql.store\_customers c ON c.cust\_id = d.cust\_id*

*WHERE category = 'Corporate'*

*GROUP BY d.cust\_id)*

*SELECT*

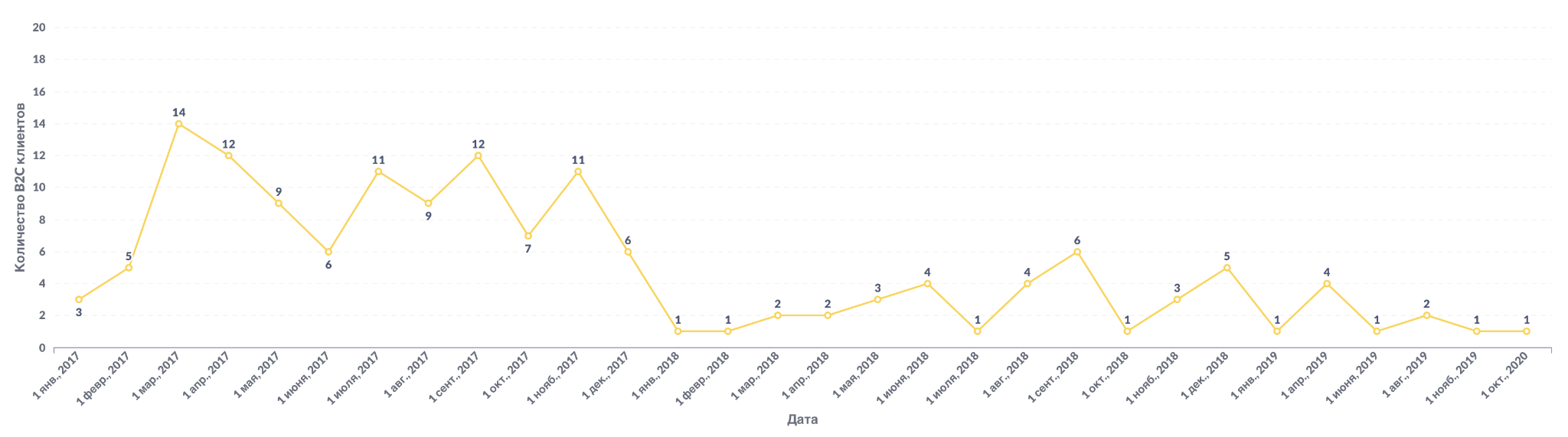
*"month",*

*COUNT(DISTINCT cust\_id) new\_custs*

*FROM first\_order\_month*

*GROUP BY "month"*

Хоть уже было выявлено, что доля B2C клиентов в общей выручки компании невелика, стоит дополнительно проверить, что доходы компании в последние годы не зависят от нее, то есть основной фокус компании не сместился с B2B клиентов. Для этого посмотрим динамику привлечения B2C клиентов и посмотрим на поведение графика. Рассматривая одинаковые промежутки времени, можно увидеть, что график имеет очень схожее поведение с предыдущим. В 2017 году количество привлеченных клиентов было стабильно выше или равно 6, но в январе 2018 года случился спад - с 6 привлеченных клиентов в декабре 2017 до 1 клиента в начале следующего года. Эта тенденция сохранилась до конца временного периода, ровно так же, как это произошло с B2B клиентами. Таким образом, можно сделать вывод, что выручка компании в целом зависит от уже привлеченных клиентов, а новых привлекается относительно малое количество. И по соотношению долей общей выручки среди двух категорий клиентов основной и приоритетной так же остается B2B.



*WITH first\_order\_month AS*

*(SELECT*

*MIN(DATE\_TRUNC('month', order\_date))::date "month",*

*d.cust\_id*

*FROM sql.store\_delivery d*

*JOIN sql.store\_customers c ON c.cust\_id = d.cust\_id*

*WHERE category = 'Consumer'*

*GROUP BY d.cust\_id)*

*SELECT*

*"month",*

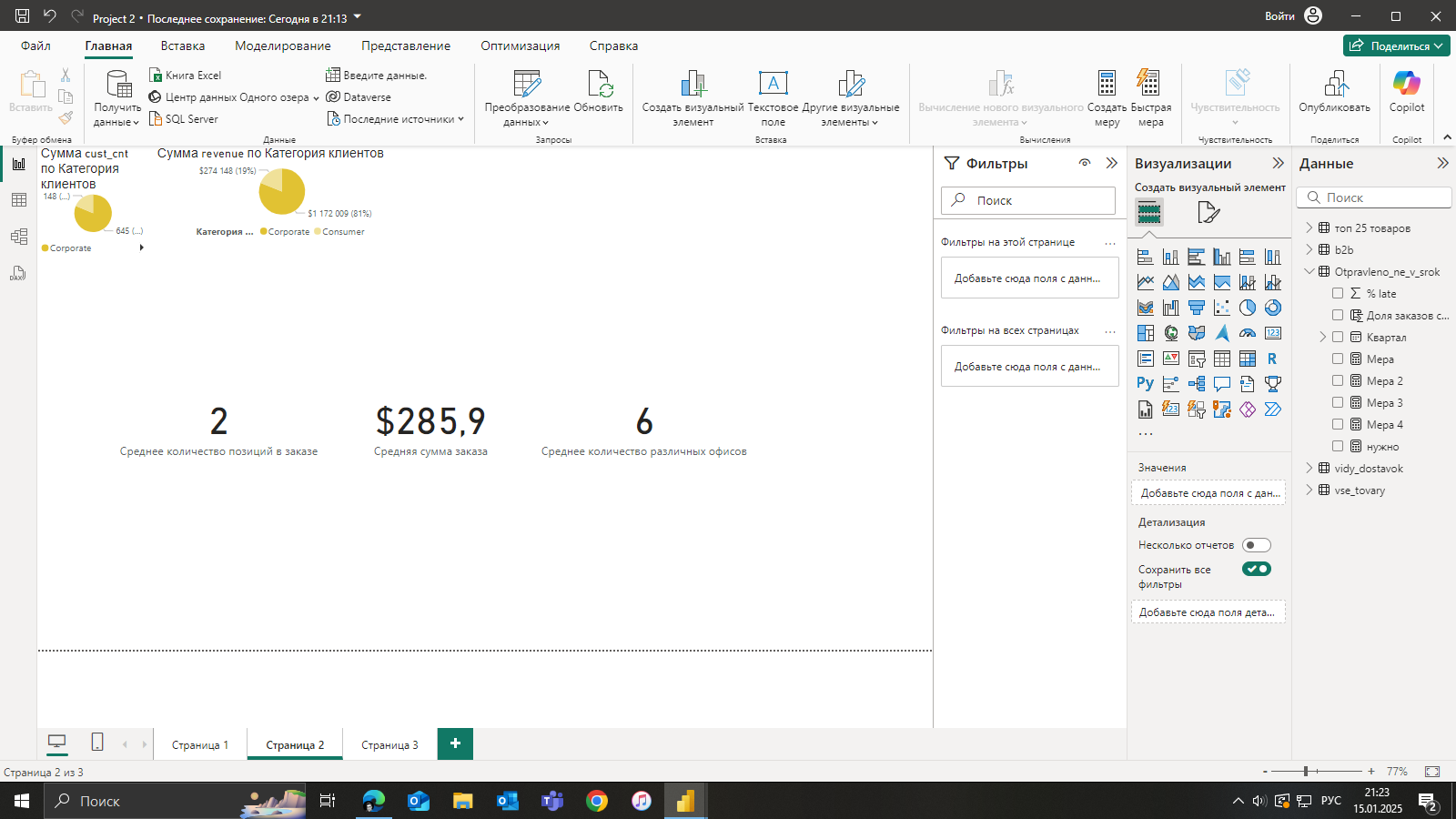
*COUNT(DISTINCT cust\_id) new\_custs*

*FROM first\_order\_month*

*GROUP BY "month"*

**Характеристика B2B клиентов:**

Убедившись, что основным направлением компании является B2B сектор, и проследив динамику привлечения новых пользователей, можно детальнее рассмотреть основные показатели по корпоративным клиентам. Важно узнать сколько в среднем у них товаров в заказах, как средняя сумма заказа и сколько в среднем различных офисов у корпоративных клиентов. Это поможет посмотреть более детально на их поведение, а также рассмотреть, насколько большими компаниями являются партнеры. В дальнейшем это может сыграть роль в выборе места открытия нового оффлайн-магазина с целью оптимизации всех бизнес процессов.



Итак, средняя сумма заказа примерно 286$ при заказе в среднем 2 различных единиц товара, что в общем и целом является не таким крупным единичным чеком, однако учитывая постоянную надобность в типах товаров, которые продает компания, это создает уверенность в долгосрочном и стабильном потоке денежных средств.

*WITH quantity AS*

*(SELECT*

*c.order\_id,*

*COUNT(c.product\_id) quantity*

*FROM*

*sql.store\_delivery d*

*JOIN sql.store\_carts c ON d.order\_id = c.order\_id*

*JOIN sql.store\_products p ON c.product\_id = p.product\_id*

*JOIN sql.store\_customers sc ON sc.cust\_id = d.cust\_id*

*WHERE sc.category = 'Corporate'*

*GROUP BY 1)*

*SELECT*

*ROUND(AVG(quantity),1) number*

*FROM quantity*

*WITH order\_sum AS*

*(SELECT*

*c.order\_id,*

*SUM(price \* quantity \* (1 - discount)) revenue*

*FROM*

*sql.store\_delivery d*

*JOIN sql.store\_carts c ON d.order\_id = c.order\_id*

*JOIN sql.store\_products p ON c.product\_id = p.product\_id*

*JOIN sql.store\_customers sc ON sc.cust\_id = d.cust\_id*

*WHERE sc.category = 'Corporate'*

*GROUP BY 1)*

*SELECT*

*ROUND(AVG (revenue),1) SUMMA*

*FROM order\_sum*

*WITH quantity\_office AS*

*(SELECT*

*d.cust\_id,*

*COUNT(DISTINCT (d.zip\_code)) quantity*

*FROM*

*sql.store\_delivery d*

*JOIN sql.store\_customers c ON c.cust\_id = d.cust\_id*

*WHERE c.category = 'Corporate'*

*GROUP BY 1)*

*SELECT*

*ROUND(AVG (quantity)) avg\_quantity\_office*

*FROM quantity\_office*

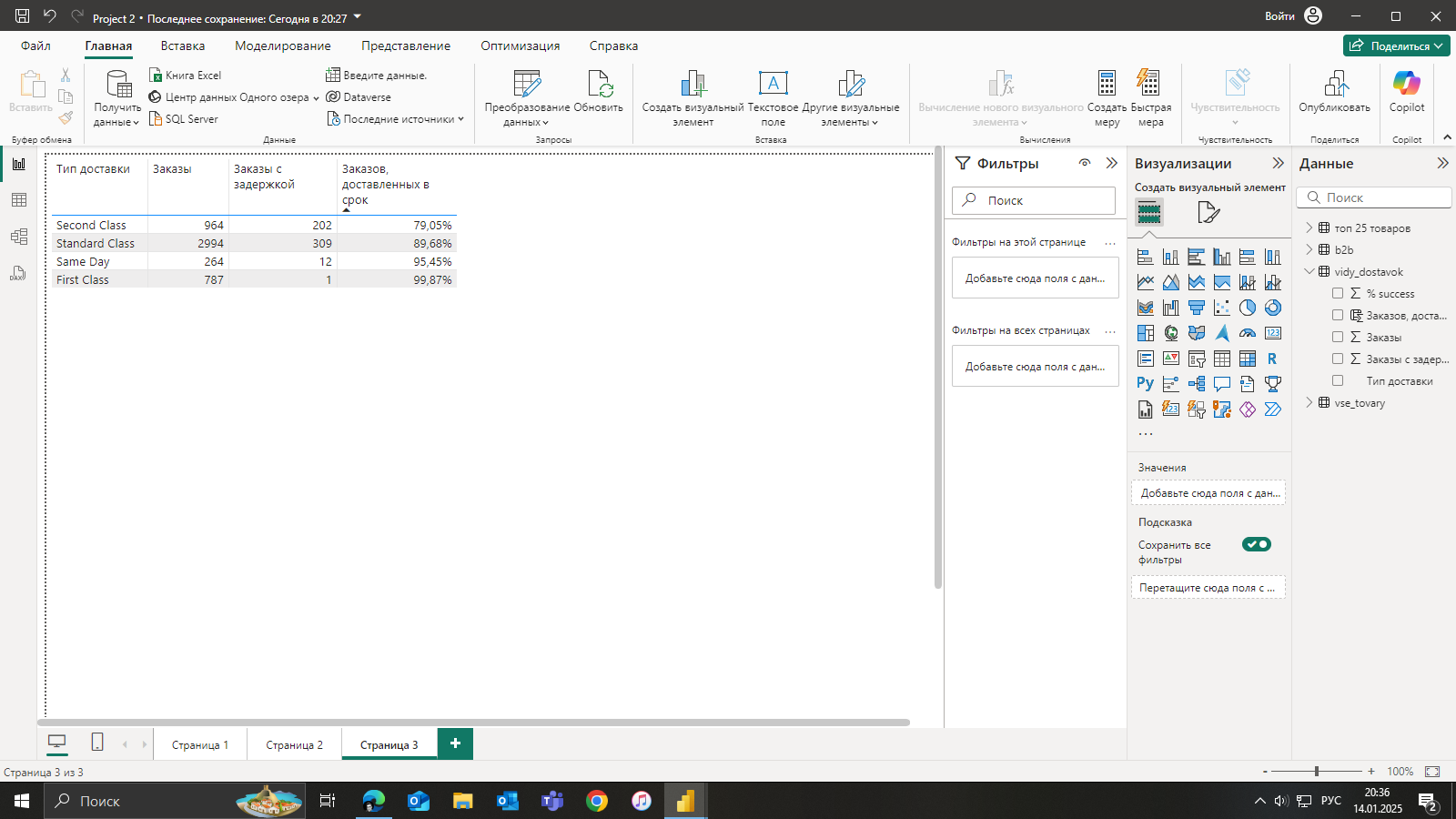
# **3) Логистика компании**

**Эффективность доставки:**

Компания Pens and Pencils оперирует онлайн. На данный момент у нее есть всего одна физическая точка - склад, откуда отправляются все товары клиентам, и находится он в Техасе, город Хьюстон. Для доставки товаров используется 4 типа доставки, перечисленных с их характеристиками в таблице: Standard Class, Second Class, First Class и Same Day.



Логистика любой компании также нуждается в постоянном контролировании и мониторинге, поэтому следует узнать, насколько доставка соответствует временным рамкам, которые заявлены для ее типа. Можно увидеть, что больше всего заказов, доставленных с задержкой были в Standard Class. Вместе с этим этот тип доставки является самым популярным - 2994 заказов были сделаны, используя именно его. Общий процент при этом успешно доставленных заказов в срок - более 89%, что не является худшим показателем. Самый маленький процент заказов, которые были доставлены в срок, имеет Second Class - 79,05%. То есть заказы, отправленные вторым классом приходят с задержкой чаще всего, на что обязательно следует обратить внимание.



*WITH delivery\_type AS*

*(SELECT*

*ship\_mode,*

*COUNT(DISTINCT sd.order\_id) orders\_cnt,*

*COUNT(CASE WHEN ship\_mode = 'Standard Class' AND (ship\_date - order\_date) > 6 THEN (ship\_date - order\_date)*

*WHEN ship\_mode = 'Second Class' AND (ship\_date - order\_date) > 4 THEN (ship\_date - order\_date)*

*WHEN ship\_mode = 'First Class' AND (ship\_date - order\_date) > 3 THEN (ship\_date - order\_date)*

*WHEN ship\_mode = 'Same Day' AND (ship\_date - order\_date) >= 1 THEN (ship\_date - order\_date) END) status*

*FROM sql.store\_delivery sd*

*GROUP BY 1)*

*SELECT*

*ship\_mode,*

*orders\_cnt,*

*status,*

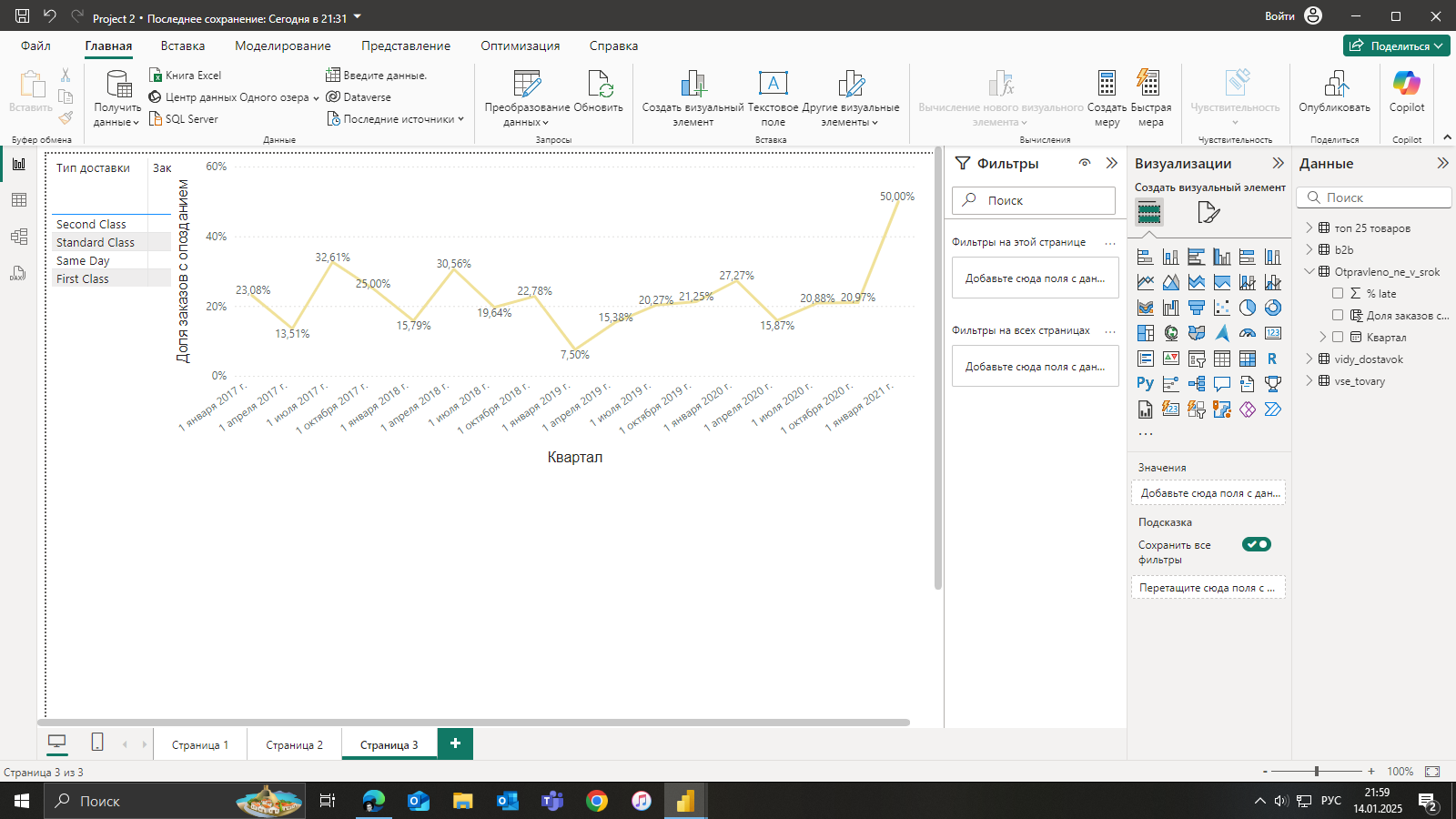
*ROUND((SUM(orders\_cnt) - SUM(status)) / SUM(orders\_cnt) \* 100, 2)::float "% success"*

*FROM delivery\_type*

*GROUP BY ship\_mode, orders\_cnt, status*

*ORDER BY "% success"*

В поисках причины этому явлению, рассмотрим насколько часто заказы, отправленные вторым классом, приходили с опозданием на всем временном промежутке. На графике видно, что довольно серьезные проблемы с доставкой наблюдались не просто на протяжении какого-то временного интервала, а каждый месяц. Это происходит систематически. Четко видно, что маленький процент задержанных заказов был лишь один раз - в январе 2019 года. Во все остальные месяцы этот класс доставки не может отличиться пунктуальностью: чаще всего задерживается более 20% заказов, в то время как в январе 2021 года этот показатель составил и вовсе огромные 50%, достигнув пикового значения. Это свидетельствует о том, что это не случайные выбросы таких значений, а какая-то проблема, недоработка, которая мешает этому типу доставки соответствовать спецификациям.



*WITH second\_class\_late AS*

*(SELECT*

*DATE\_TRUNC('quarter',ship\_date)::date quarter,*

*ship\_mode,*

*order\_id,*

*CASE WHEN ship\_mode = 'Second Class' and ship\_date - order\_date > 4 THEN 'late' END late\_orders*

*FROM sql.store\_delivery*

*WHERE ship\_mode = 'Second Class')*

*SELECT*

*quarter,*

*ROUND(100 \* COUNT(late\_orders)/COUNT(order\_id)::numeric, 2) "% late"*

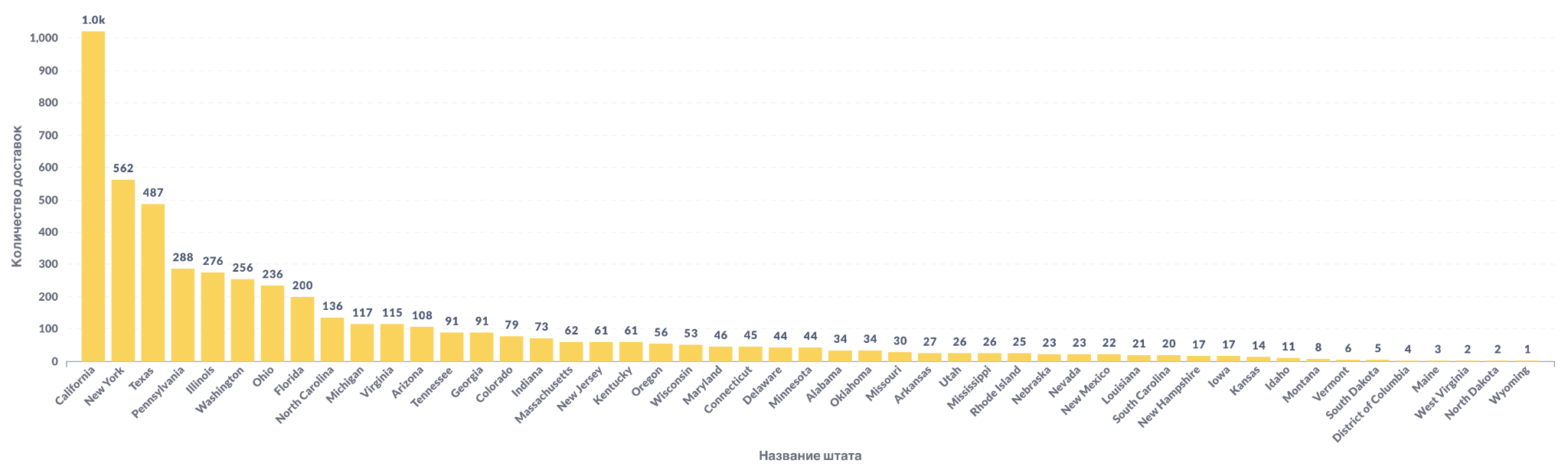
*FROM second\_class\_late*

*GROUP BY 1*

*ORDER BY 1*

# **4) Выбор офлайн точки продаж**

Частично эту проблему можно закрыть в том числе и открытием офлайн точки. Открыв свой первый офлайн магазин, компания могла бы существенно сократить издержки на доставку товаров в целом. По крайней мере на существенную долю, если найти место, откуда заказывают чаще и больше всего, и расположить магазин именно там, расширив свое присутствие географически. Для того, чтобы обнаружить такое место, построим графики с распределением заказов и выручки по штатам и городам, так как один город в штате может перетягивать на себя большинство заказов, задавая высокий уровень доходности, а может быть и ситуация, в которой множество городов распределили между собой выручку, но суммарно делая этот штат более привлекательным. На графике по распределению заказов среди штатов отчетливо видно, что Калифорния (California) занимает первое место с большим запасом - со вторым местом ее отделяет разница в 2 раза: 1000 против 562 заказов в штате Нью-Йорк (New York).



*SELECT*

*state,*

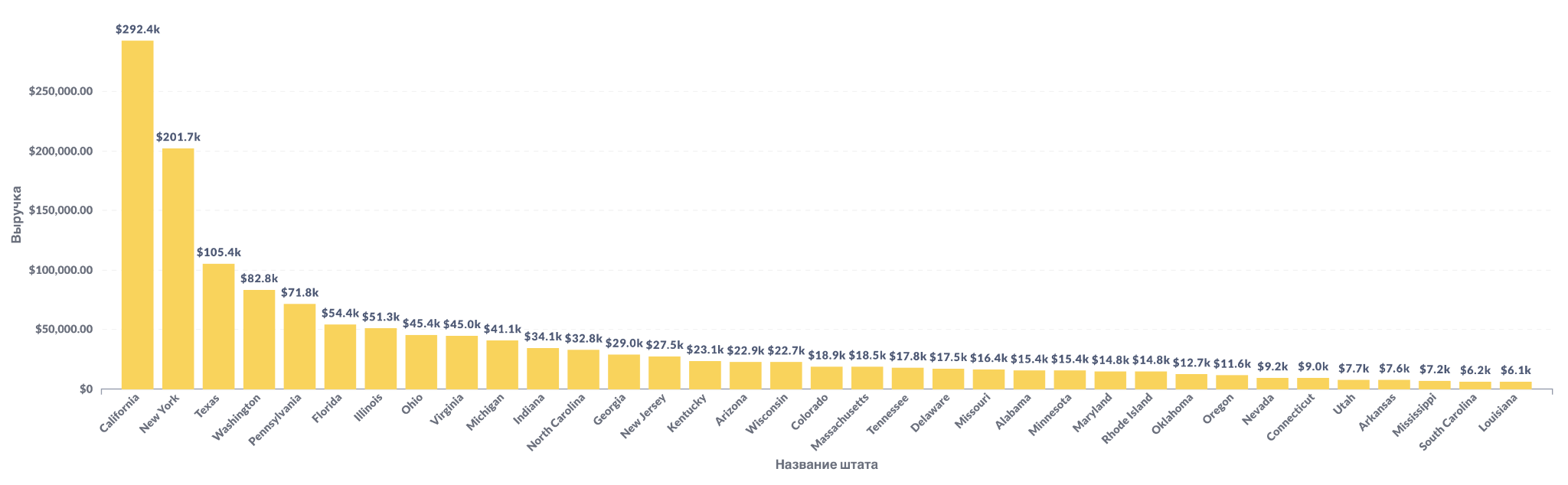
*count (order\_id) number*

*FROM sql.store\_delivery*

*GROUP BY state*

*ORDER BY number DESC*

Анализируя выручку, Калифорния также держит лидирующую позицию при значении в 292 000$, за ней идут Нью-Йорк и Техас с выручкой в 201 000$ и 105 000$ соответственно.



*SELECT*

*state,*

*ROUND(SUM((p.price \* ca.quantity) \* (1 - ca.discount))) revenue*

*FROM sql.store\_delivery sd*

*JOIN sql.store\_carts ca ON sd.order\_id = ca.order\_id*

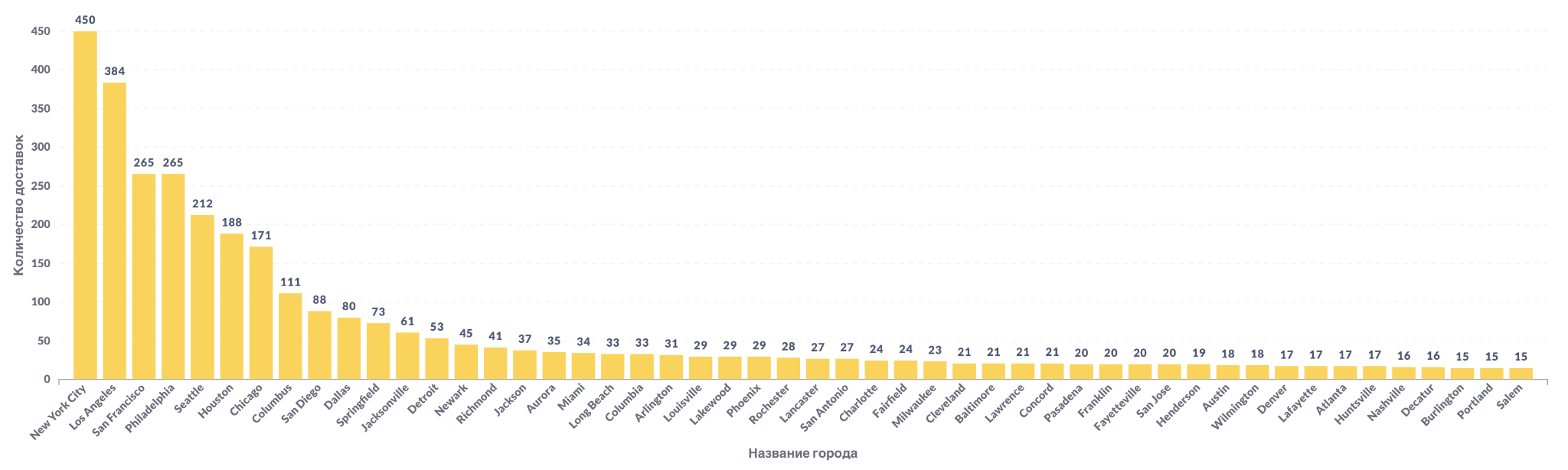
*JOIN sql.store\_products p ON ca.product\_id = p.product\_id*

*GROUP BY state*

*ORDER BY revenue DESC*

*LIMIT 35*

Таким образом, наиболее привлекательным штатом для открытия офлайн точки на основе количества заказов и выручки с них является Калифорния. Как и говорилось ранее, следует также рассмотреть и распределение заказов по городам США, так как выручка по городам штата может быть довольно распылена, а открыть магазин прямо в самом городе, где по статистике происходит больше всего заказов при прочих равных, тоже является вариантом для рассмотрения. Исходя из графика, можно понять, что больше всего заказов приходится на город Нью-Йорк, самый большой и густонаселенный город США, который находится в одноименном штате Нью-Йорк, входящий в топ-3 по заказам в целом. За ним идут Лос-Анджелес (Los Angeles) и Сан-Франциско (San-Francisco), которые расположены в штате Калифорния, а он уже возглавляет список штатов по заказам в целом.



*SELECT*

*city,*

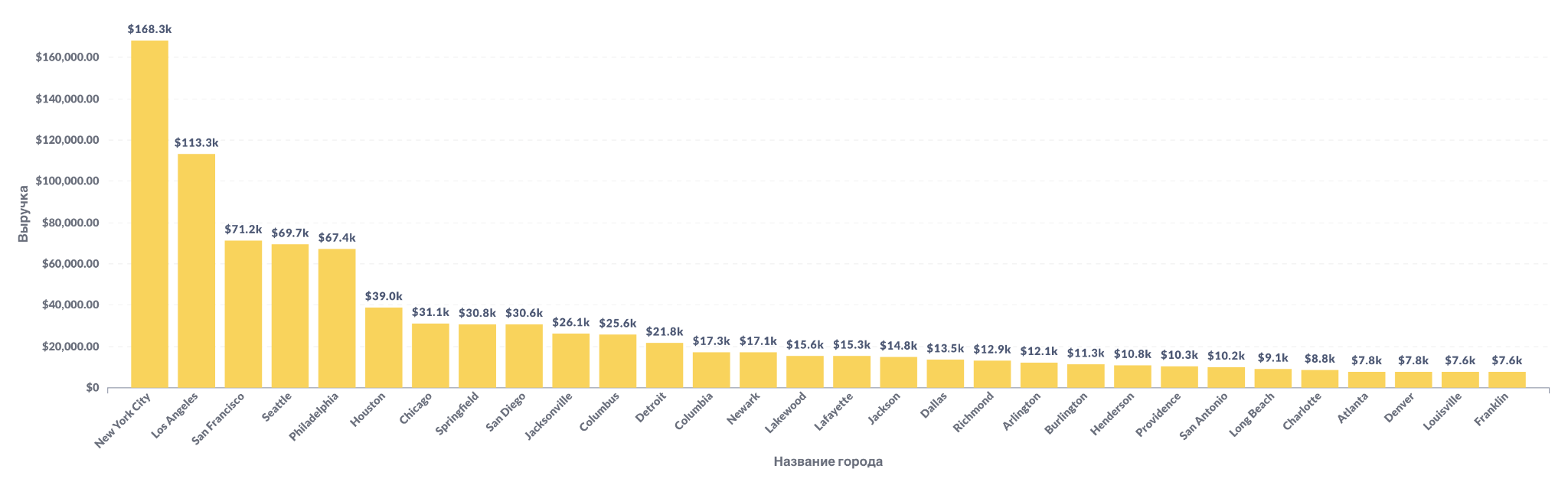
*COUNT (DISTINCT order\_id) number*

*FROM sql.store\_delivery*

*GROUP BY city*

*ORDER BY number DESC*

Нью-Йорк снова занимает первое место, за ним так же идут Лос-Анджелес и Сан-Франциско, при этом их суммарная выручка немногим больше одного лишь Нью-Йорка. То есть, если рассматривать город открытия, то этим городом точно должен стать Нью-Йорк.



*SELECT*

*city,*

*ROUND(SUM((p.price \* ca.quantity) \* (1 - ca.discount))) revenue*

*FROM sql.store\_delivery sd*

*JOIN sql.store\_carts ca ON sd.order\_id = ca.order\_id*

*JOIN sql.store\_products p ON ca.product\_id = p.product\_id*

*GROUP BY city*

*ORDER BY revenue DESC*

*LIMIT 25*

# **5) Выводы и рекомендации**

Проведя общий анализ положения дел компании, можно сделать выводы о ее особенностях ведения бизнеса, точках роста, клиентах и вариантах местоположения, в котором можно открыть офлайн-магазин:

1. В целом выручка компании стабильно растет на всем протяжении ее существования, каждый год компании удавалось увеличивать показатели эффективности и повысить выручку.
2. Среди трех категорий товаров, которые продает компания, наибольшая часть выручки приходится на Technology, что обуславливает большое количество позиций техники в топ-25 самых продаваемых товаров, что составляет 1,38% от общего количества товаров, продаваемых компанией - 1818. Этот топ-25 генерирует 18% всей выручки компании - за стабильностью производства и доставкой этих товаров следует вести пристальное наблюдение, так как значения общей выручки могут сильно варьироваться от успешности продаж этих товаров.
3. Компания больше ориентирована на B2B сектор: количество таких клиентов - 645 человек (81,34%). При этом соотношение долей выручки между B2B и B2C остается примерно таким же - 80% приходится именно на корпоративных клиентов.
4. Доход компании держится не просто за счет B2B клиентов, а за счет тех, которые были привлечены до начала 2018 года. После этого в привлечении новых клиентов случился огромный спад вплоть до 2021 года и роста уже не произошло. Точно такая же динамика прослеживается и с B2C клиентами
5. Средний чек с одного заказа у корпоративного клиента - 285,9$ при наличии двух различных единиц товара в заказе. Сам по себе чек небольшой, но учитывая категории товаров, которые так или иначе постоянно нужны в бизнес сфере, их постоянный заказ и обеспечивает стабильно растущую выручку компании.
6. Компания имеет один физический склад, но торговля проходит онлайн, поэтому используется доставка четырех типов с различными планируемыми сроками доставки. Худший показатель доставленных в срок товаров имеет тип доставки Second Class. В некоторые месяцы до заказчиков не доходило вовремя 50% заказов. Это свидетельствует о систематических проблемах, возникающих в процессах доставки. Это второй тип доставки по популярности и такое частое отклонение сроков доставки в отрицательную сторону может негативно повлиять на лояльность клиентов.
7. Для открытия офлайн точки есть несколько вариантов развития событий. Если рассматривать город, в котором открыть магазин с целью максимального снижения количества доставок и логистических издержек, то максимально под эту задачу подходит город Нью-Йорк из одноименного штата - на этот город приходится самое большое количество заказов и самая большая доля выручки. К тому же, это самый густонаселенный и большой город в США.
8. Также есть вариант открытия магазина в самом популярном по заказам штате - Калифорнии. Ее города Лос-Анджелес и Сан-Франциско занимают 2е и 3е место соответственно после Нью-Йорка. В целом на этот штат приходится самое большое количество заказов и выручки, поступающих с них. Но, это не сократит издержки на логистику максимально как в случае с открытием в городе Нью-Йорк. Это позволит их снизить, но не лучшим образом, потому что такое большое количество заказов всего штата сильно раздроблено между городами этого штата, но зато позволит находиться в самом популярном по заказам штате США.