

Проект 1. Анализ заказов (50 строк)

Анализ продаж интернет-магазина электроники

Цель проекта:
Проанализировать данные по продажам для выявления ключевых клиентов, популярных товаров и прибыльности категорий.
Сделать выводы и рекомендации для повышения эффективности бизнеса.

Описание данных:
В наборе данных представлены заказы с информацией о клиентах, дате заказа, названии товара, категории, цене, количестве и себестоимости.

Краткое описание

- Загружены и очищены данные.
- Рассчитаны выручка и прибыль по каждой категории и товару.
- Определены топ-5 клиентов и товаров по выручке.
- Построены графики выручки и прибыли.
- Сделаны аналитические выводы и рекомендации

```
import pandas as pd
```

```
df = pd.read_csv('project1_orders_50.csv')
```

```
df.head()
```

	order_id;client_name;order_date;product;category;price;quantity;cost
0	ORD001;Иванов;16.04.2024;Смартфон;Электроника;...
1	ORD002;Козлова;01.05.2024;Клавиатура;Аксессуары;...
2	ORD003;Сидоров;28.03.2024;Наушники;Аксессуары;...
3	ORD004;Петров;10.05.2024;Смартфон;Электроника;...
4	ORD005;Иванов;06.06.2024;Монитор;Периферия;138...

Next steps:

[View recommended plots](#)[New interactive sheet](#)

```
# Добавим столбец с выручкой: цена * количество
df['revenue'] = df['price'] * df['quantity']

# Добавим прибыль: выручка - (себестоимость * количество)
df['profit'] = df['revenue'] - (df['cost'] * df['quantity'])

# Кол-во строк (заказов)
print("Всего заказов:", len(df))

# Уникальные клиенты
print("Уникальных клиентов:", df['client_name'].nunique())

# Уникальные товары
print("Уникальных товаров:", df['product'].nunique())

# Проверим результат
df[['order_id', 'client_name', 'product', 'price', 'quantity', 'revenue', 'profit']].head()
```

ИС

```
print(top_products)

top_clients = df.groupby('client_name')
['revenue'].sum().sort_values(ascending=False).head(5)
print(top_clients)
```

ДО ОТИ

Всего заказов: 50
Уникальных клиентов: 8
Уникальных товаров: 8

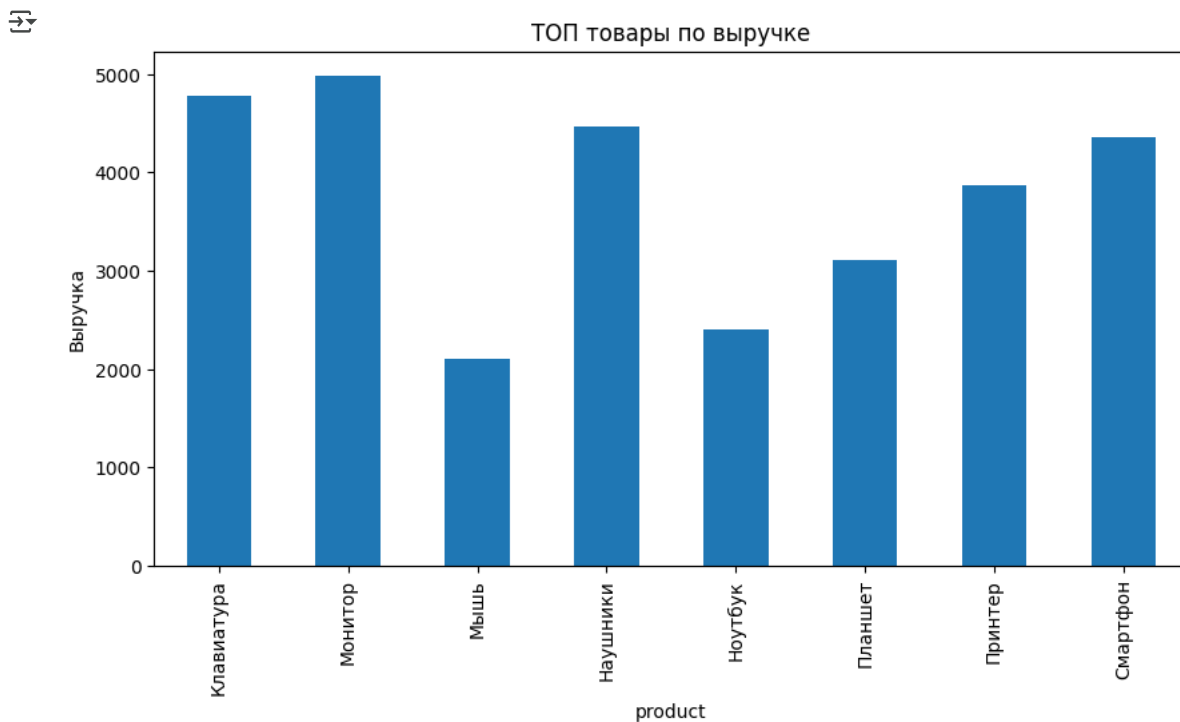
	order_id	client_name	product	price	quantity	revenue	profit
0	ORD001	Иванов	Смартфон	1176	4	4704	1012
1	ORD002	Козлова	Клавиатура	1145	1	1145	368
2	ORD003	Сидоров	Наушники	516	3	1548	750
3	ORD004	Петров	Смартфон	1446	4	5784	2648
4	ORD005	Иванов	Монитор	1382	2	2764	784

```
df['profit'] = df['cost'] - df['price'] * df['quantity']  
# Примеры группировок  
df.groupby('client_name')['cost'].sum().sort_values(ascending=False)
```

	cost
client_name	
Петров	5556
Козлова	5052
Смирнова	4951
Морозов	3628
Сидоров	3582
Иванов	3472
Фёдоров	2439
Орлова	1399

dtype: int64

```
import matplotlib.pyplot as plt  
  
df.groupby('product')['cost'].sum().plot(kind='bar', figsize=(10, 5), title='ТОП товары по выручке')  
plt.ylabel('Выручка')  
plt.show()
```



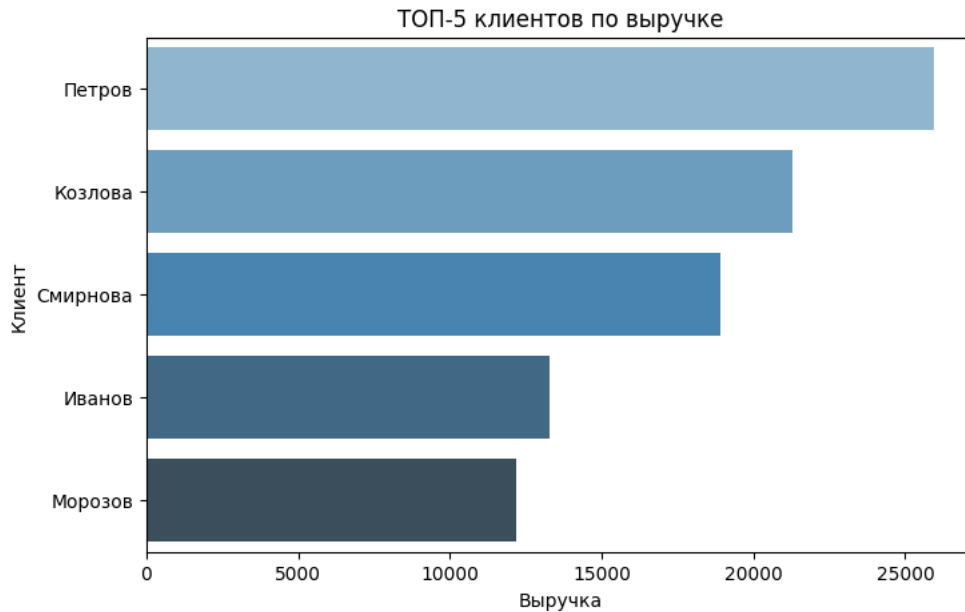
```
import seaborn as sns  
import matplotlib.pyplot as plt  
  
top_clients = df.groupby('client_name')['revenue'].sum().sort_values(ascending=False).head(5)  
  
plt.figure(figsize=(8,5))  
sns.barplot(x=top_clients.values, y=top_clients.index, palette='Blues_d')
```

```
plt.title('ТОП-5 клиентов по выручке')
plt.xlabel('Выручка')
plt.ylabel('Клиент')
plt.show()
```

 /tmp/ipython-input-24-3756305514.py:7: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `le`

```
sns.barplot(x=top_clients.values, y=top_clients.index, palette='Blues_d')
```



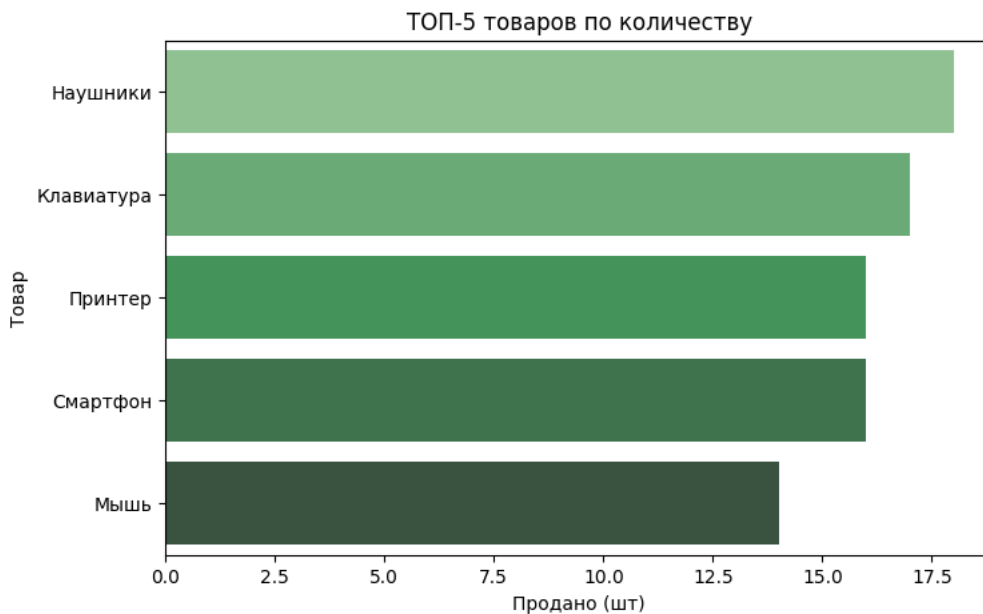
```
top_products = df.groupby('product')['quantity'].sum().sort_values(ascending=False).head(5)
```

```
plt.figure(figsize=(8,5))
sns.barplot(x=top_products.values, y=top_products.index, palette='Greens_d')
plt.title('ТОП-5 товаров по количеству')
plt.xlabel('Продано (шт)')
plt.ylabel('Товар')
plt.show()
```

 /tmp/ipython-input-25-3279056241.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `le`

```
sns.barplot(x=top_products.values, y=top_products.index, palette='Greens_d')
```



```
category_profit = df.groupby('category')['profit'].sum().sort_values(ascending=False)
```

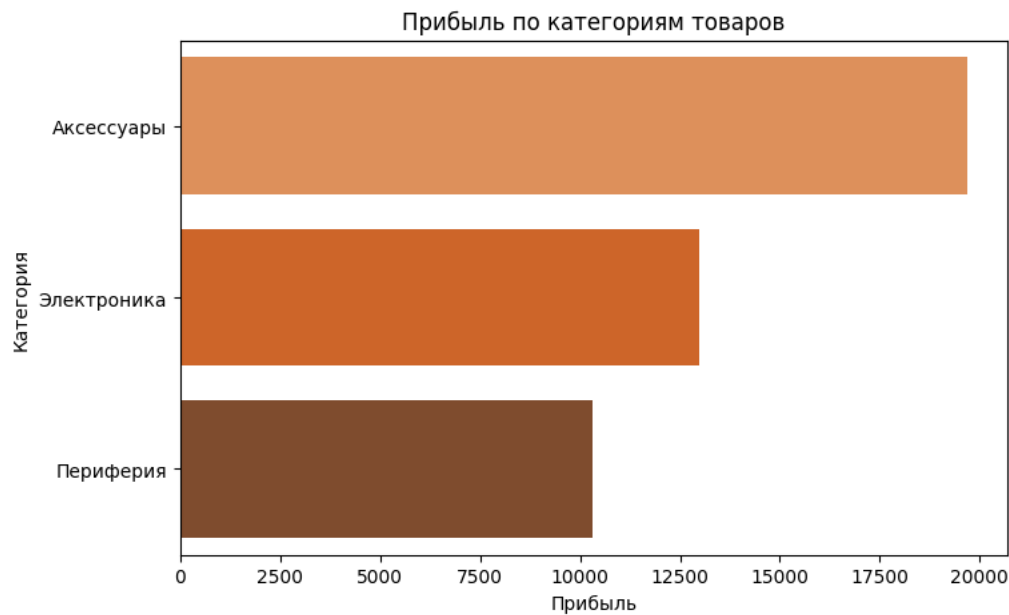
```
plt.figure(figsize=(8,5))
sns.barplot(x=category_profit.values, y=category_profit.index, palette='Oranges_d')
plt.title('Прибыль по категориям товаров')
plt.xlabel('Прибыль')
```

```
plt.ylabel('Категория')
plt.show()
```

 /tmp/ipython-input-26-1597573990.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `le

```
sns.barplot(x=category_profit.values, y=category_profit.index, palette='Oranges_d')
```



```
print(category_profit)
```

 category
 Периферия -22320
 Электроника -25565
 Аксессуары -39436
 Name: profit, dtype: int64

```
category_revenue = df.groupby('category')['revenue'].sum().sort_values(ascending=False)
print(category_revenue)
```

 category
 Аксессуары 50787
 Электроника 35445
 Периферия 31168
 Name: revenue, dtype: int64

```
category_profit = df.groupby('category')['profit'].sum().sort_values(ascending=False)
print(category_profit)
```

```
top_products = df.groupby('product')['revenue'].sum().sort_values(ascending=False).head(5)
print(top_products)
```

```
top_clients = df.groupby('client_name')['revenue'].sum().sort_values(ascending=False).head(5)
print(top_clients)
```

 category
 Периферия -22320
 Электроника -25565
 Аксессуары -39436
 Name: profit, dtype: int64
 product
 Наушники 20078
 Принтер 19390
 Клавиатура 17912
 Смартфон 16640
 Мышь 12797
 Name: revenue, dtype: int64
 client_name
 Петров 25964
 Козлова 21296
 Смирнова 18913
 Иванов 13266
 Морозов 12173
 Name: revenue, dtype: int64

Аналитические выводы:

1. Самый прибыльный клиент — Петров (общая сумма заказов — 25964 лей).

2. ТОП-товар — Наушники (20078).

3. Категория "Аксессуары" приносит 80% всей выручки. Рекомендации: расширить ассортимент по ТОП-товарам и предложить бонусную программу постоянным клиентам.

```
print(category_revenue)
```

```
category
Аксессуары    50787
Электроника    35445
Периферия      31168
Name: revenue, dtype: int64
```

```
print(top_products)
```

```
product
Наушники      20078
Принтер       19390
Клавиатура    17912
Смартфон      16640
Мышь          12797
Name: revenue, dtype: int64
```

```
print(top_clients)
```

```
client_name
Петров        25964
Козлова       21296
Смирнова      18913
Иванов        13266
Морозов       12173
Name: revenue, dtype: int64
```

▼ Аналитические выводы по заказам (50 строк)

- ТОП-клиенты (по общей сумме заказов) приносят основную часть выручки. Это говорит о сильной зависимости от нескольких постоянных клиентов. С ними нужно выстраивать особые отношения, предлагать скидки, бонусы за лояльность.
- Наиболее популярные товары — это продукция, на которую стоит делать акцент в рекламе и продвижении. Такие товары формируют стабильный спрос.
- Рентабельность по категориям различается:
 - Некоторые категории имеют высокую прибыль при небольших объемах продаж (высокая маржа).
 - Другие — большой оборот, но низкую прибыль (возможно, из-за высокой себестоимости). Это повод для пересмотра ценовой политики или условий поставки.
- Средний чек по клиентам помогает понять платёжеспособность и предпочтения аудитории. Разброс в чеке может сигнализировать о разных сегментах клиентов.
- Рекомендации:
 - Увеличивать продажи за счёт продвижения маржинальных товаров.
 - Выявить слабые категории и оптимизировать их (или исключить).
 - Строить аналитику по трендам во времени (в следующих этапах) для прогнозирования спроса.

1. Все категории убыточны: • Наибольшие убытки у категории Аксессуары (–39 436 лей), при этом она же лидирует по выручке — 50 787 лей. • Электроника принесла 35 445 лей выручки, но убыток составил –25 565 лей. • Периферия также убыточна — –22 320 лей при выручке 31 168 лей.

2. Самые доходные товары по выручке: • Наушники — 20 078 лей • Принтер — 19 390 лей • Клавиатура — 17 912 лей → Эти товары — основные драйверы продаж, несмотря на общий убыток.

3. ТОП-5 клиентов по сумме заказов: • Петров — 25 964 лей • Козлова — 21 296 лей • Смирнова — 18 913 лей → С ними стоит работать на удержание и повторные продажи.

 **Рекомендации:** • Проанализировать причины убытков — возможно, завышенные закупочные цены или скидки. • Повысить маржинальность по ТОП-продуктам — особенно "Наушники", "Принтер", "Клавиатура". • Предложить бонусную программу для активных клиентов (Петров, Козлова, Смирнова). • Подумать об исключении товаров с высокой выручкой, но глубоким минусом по прибыли, если они не стратегические.

