**АНАЛІТИЧНИЙ ЗВІТ Olist**

**Про набір даних**

**Бразильський набір публічних даних електронної комерції від Olist**

Це публічний набір даних про замовлення електронної комерції в Бразилії, зроблені в [магазині Olist](http://www.olist.com/) .

Набір даних містить інформацію про 100 тисяч замовлень з 2016 по 2018 рік, зроблених на кількох торгових майданчиках Бразилії. Данні в датасеті дозволяють переглядати замовлення з різних точок зору: від статусу замовлення, ціни, оплати та ефективності доставки до місцезнаходження клієнта, атрибутів продукту та, нарешті, відгуків, написаних клієнтами, а також містить дані геолокацій, які пов'язують поштові індекси Бразилії з географічними координатами.

*Повний набір даних Olist, що охоплює всі ключові аспекти онлайн-торгівлі: замовлення, клієнти, продавці, геолокація, платежі, огляди, товари та категорії. (складається з 9 таблиць в форматі CSV)*

**Всього ми маємо:**

* **99 441 клієнтів**
* **1 000 163 записів геолокації**
* **112 650 товарних позицій у замовленнях**
* **103 886 платежів**
* **99 224 оглядів**
* І аналогічно велика кількість даних по замовленнях, продавцях і продуктах.

**Методи аналізу:**

За допомогою Python проведено об'єднання всіх таблиць та проведено основний аналіз даних з їх візуалізацією.

Основну візуалізацію оформлено в **Power Bi**

Для повноти аналізу мною вибрано **9 аналітичних напрямків, які можна реалізувати на основі злитої таблиці, а саме:**

**АНАЛІЗ №1: Динаміка продажів у часі**

**Цілі:**

* Дослідити кількість замовлень по місяцях;
* Аналіз доходу (оплати) по місяцях;
* Побачити тренди зростання/спаду активності;
* Виявити сезонні піки.

**АНАЛІЗ №2: Географічний аналіз клієнтів і продавців**

**Дослідити:**

* Топ-штати та міста за кількістю замовлень
* Географічний розрив між клієнтом і продавцем
* Основні центри комерції

**АНАЛІЗ №3: Популярність категорій товарів**

**Дослідимо:**

* Топ-10 категорій по кількості продажів
* Прибутковість категорій
* Середній чек по категоріях

**АНАЛІЗ №4: Оцінки та задоволеність клієнтів**

**Дослідити:**

Розподіл оцінок

Зв’язок між оцінками та категоріями

Регіональні відмінності

**АНАЛІЗ №5: Час доставки та очікування**

**Мета:**

порівняти фактичну та очікувану доставку, виявити затримки.

**АНАЛІЗ №6: Структура платежів та середній чек**

**Дослідити:**

Розподіл типів оплат

Середній чек по типах платежів

**АНАЛІЗ №7: Продавці**

**Цілі:**

Рейтинг продавців за кількістю замовлень

Середній чек по продавцях

Рейтинг на основі оцінок

**АНАЛІЗ №8: Вартість доставки та задоволення**

**Ціль:**

Чи впливає сума доставки на оцінку клієнтів?

**АНАЛІЗ №9: Випадки затримки доставки**

**Ціль:**

Визначити частку замовлень із затримкою

Виявити категорії та регіони, де це відбувається найчастіше

**На підставі отриманих даних та візуалізацій можу зробити висновки:**

## 1. Динаміка продажів у часі

* Найвища активність спостерігається з кінця 2017 до середини 2018 року.
* Сезонні піки замовлень — весна та передноворічний період.
* Доцільно інвестувати в маркетинг у 'тихі' місяці (січень–лютий).

## 2. Географічний аналіз клієнтів і продавців

* Найбільше замовлень із штатів SP, RJ, MG.
* Основні хаби продавців: São Paulo, Belo Horizonte, Curitiba.
* Варто оптимізувати логістику в південно-східному регіоні.

## 3. Популярність категорій товарів

* Найчастіше купують: 'bed\_bath\_table', 'health\_beauty', 'sports\_leisure'.
* Найвищий чек у категоріях: 'computers\_accessories', 'furniture\_decor'.
* Необхідно стимулювати попит у прибуткових, але не популярних категоріях.

## 4. Рівень задоволеності клієнтів

* 57% клієнтів ставлять 5 зірок.
* Високий рейтинг у категоріях книг, музики, телеком.
* Варто покращити обслуговування у великогабаритних категоріях.

## 5. Час доставки

* Переважна більшість доставок прибуває вчасно або раніше.
* Затримки до 10 днів часті у складних категоріях (меблі, кондиціонери).

## 6. Структура платежів

* 75% покупок оплачуються кредитною карткою.
* Найвищий середній чек — банківський переказ та ваучери.
* Вигідно стимулювати оплати через 'boleto' у преміум-сегменті.

## 7. Аналіз продавців

* Найактивніші продавці мають >100 замовлень.
* Середній чек та рейтинг варіюється між продавцями.
* Можна впровадити рейтингову систему підтримки топ-продавців.

## 8. Вартість доставки vs оцінка

* Вищі витрати на доставку пов’язані з нижчими оцінками (1–2 зірки).
* Важливо або знижувати ціни, або компенсувати їхні враження.

## 9. Затримки в доставці

* 7.3% замовлень затримані.
* Найгірші результати в категоріях: 'furniture\_bedroom', 'air\_conditioning'.
* Потрібна спеціальна політика SLA для цих товарів.

## Загальні бізнес-рекомендації:

- Автоматизація логістичного SLA контролю.  
- Персоналізовані маркетингові кампанії за кластерами.  
- Зниження ризику відтоку шляхом аналізу оцінок і затримок.  
- Запуск програм лояльності для продавців і VIP-клієнтів.  
- Оптимізація цінової політики та структури доставок.

**Посилання:**

**код Python:**

<https://colab.research.google.com/drive/1jWSn4555NQK13iLTSBhAEOU-Fii-oF5M?usp=sharing>

<https://github.com/Khanbir/Data_analitycs/blob/main/Final_projekt_Khanbir_2.ipynb>

**дашборд:**

<https://drive.google.com/file/d/1tMkZTzr-tiwaaAXVg2HXPABZyPqa2Weo/view?usp=sharing>

<https://drive.google.com/file/d/14fMl2DCCSw-haLmn2X8m0djHYvixNblp/view?usp=sharing>