



---

# HURTOWNIE DANYCH

---

Projekt – Analiza danych platformy e-commerce Olist (Brazylia)



ALEKSANDER STEPANIUK

NR. INDEKSU: 272644

Politechnika Wrocławska, Informatyka Stosowana

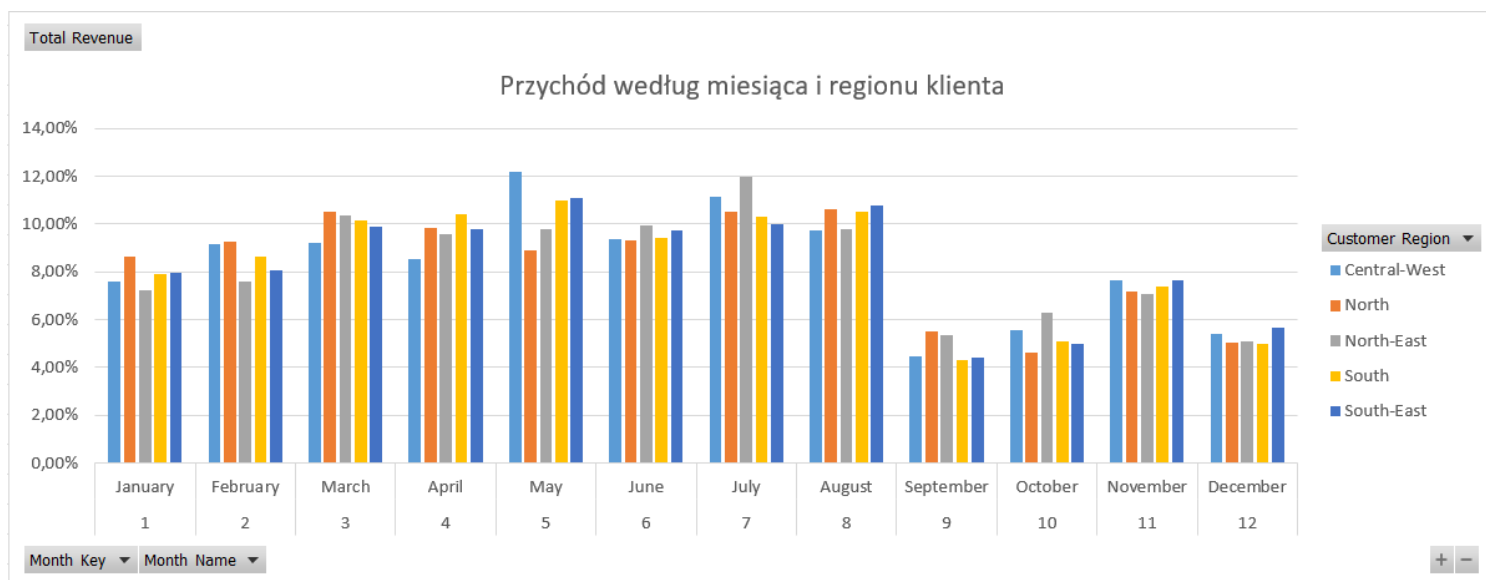
## Etap 3 – 02.06.2025 r.

Kostka:

1. TimeDim – wymiar czasu
  - Rok, kwartał, miesiąc, dzień tygodnia, godzina
2. PaymentDim
  - Typ płatności
3. ProductDim
  - Kategoria produktu
4. SellerDim
  - Miasto, stan, region, gęstość zaludnienia miasta, powierzchnia, populacja
5. CustomerDim
  - Miasto, stan, region, gęstość zaludnienia miasta, powierzchnia, populacja, tag klienta mówiący czy jest stałym czy powracającym klientem
6. ReviewDim
  - Gwiazdkowa ocena zamówienia
7. FactOrders
  - payment\_value – wartość całego zamówienia
  - total\_revenue – wartość przedmiotu + koszty przesyłki
  - total\_items – ilość przedmiotów
  - review\_score – ilość gwiazdek zostawiona w ocenie zamówienia od 0 do 5
  - delivery\_time – liczone w dniach od momentu zamówienia do dostawy
  - average\_review\_score – miara kalkulowana licząca średnią ilość gwiazdek
  - average\_delivery\_time – miara kalkulowana do liczenia średniego czasu dostawy
  - maximum\_payment\_value – miara kalkulowana, maksymalna wartość całego zamówienia

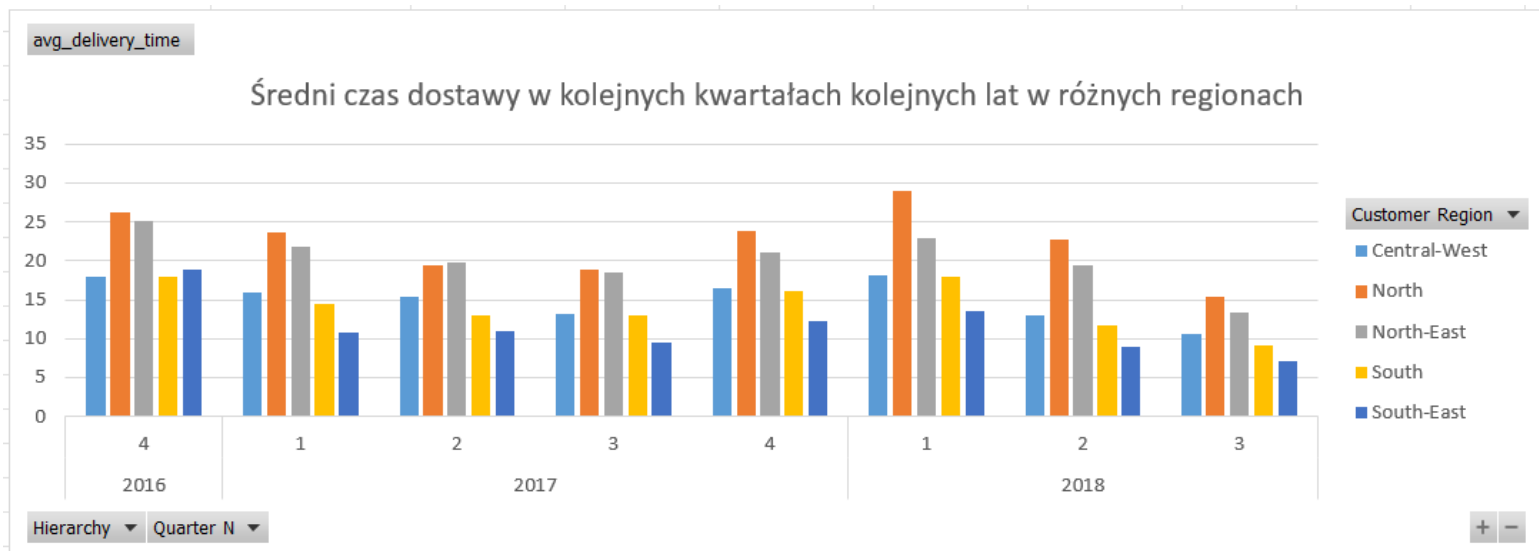
Zestawienia:

### 1. Procentowy przychód według miesiąca i regionu klienta



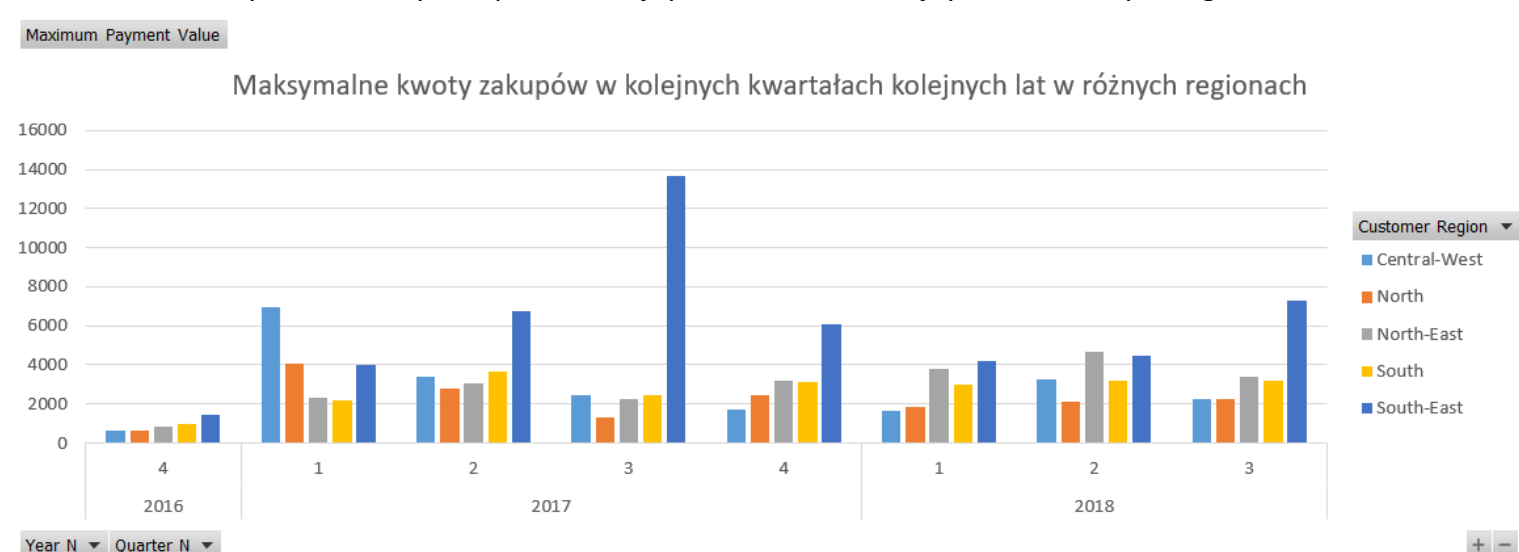
Widać większą sprzedaż w miesiącach od stycznia do sierpnia, niezależnie od obszaru (na co nie ma wpływu ograniczony zbiór danych, ponieważ dane zaczynają się od października 2016, a trwają do września 2018).

## 2. Średni czas dostawy w kolejnych kwartałach kolejnych lat w różnych regionach



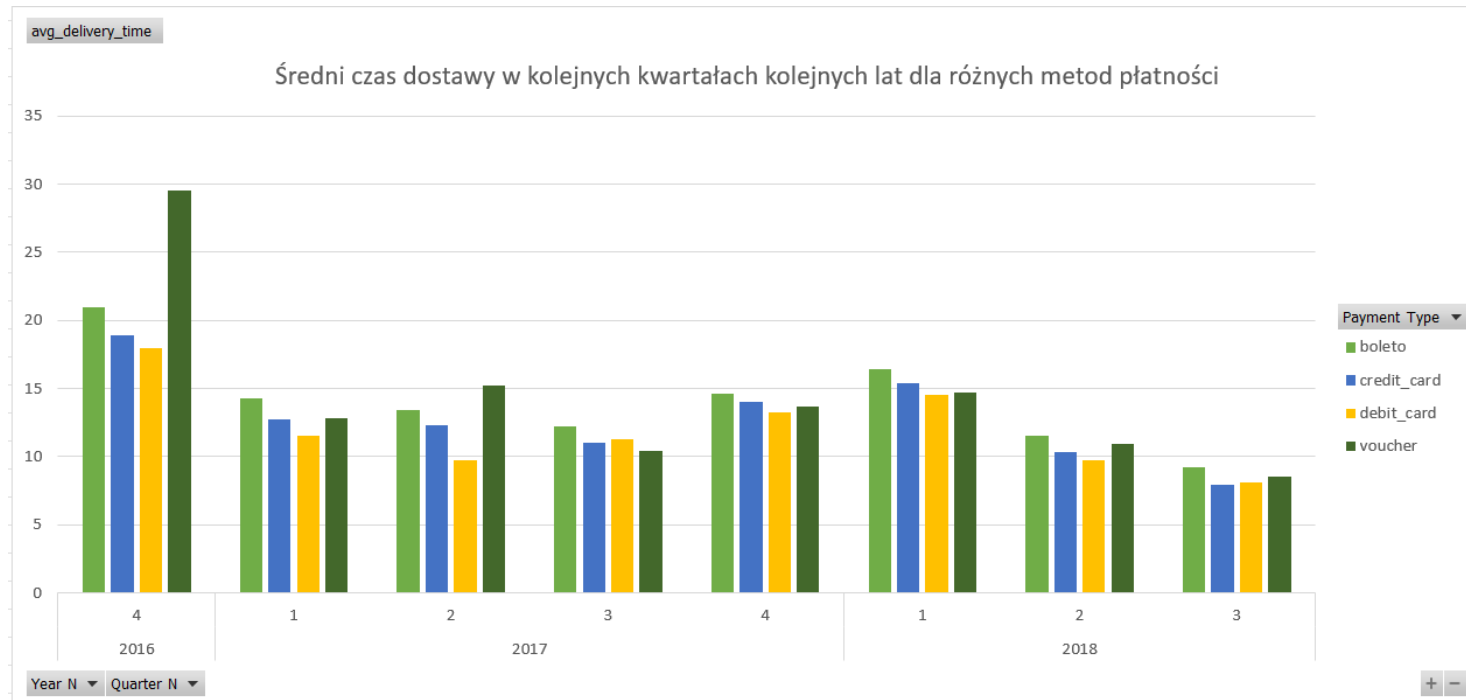
Widać że w regionie północnym i północno-wschodnim czas dostawy jest znacznie wyższy od pozostałych regionów, niezależnie od kwartału i roku są to zawsze regiony z najdłuższym czasem dostawy (około tydzień dłużej niż w innych regionach) - może to być spowodowane mniejszym zaludnieniem tego regionu oraz mniejszą sprzedażą w nim. Z drugiej strony region południowo-wschodni gdzie znajdują się największe miasta brazylii takie jak Brasilia czy Sao Paulo mają zdecydowanie najkrótszy czas dostawy. Natomiast widać także, że wraz z biegiem czasu ogólny czas dostawy ma tendencje spadkową.

## 3. Maksymalne kwoty zakupów w kolejnych kwartałach kolejnych lat w różnych regionach



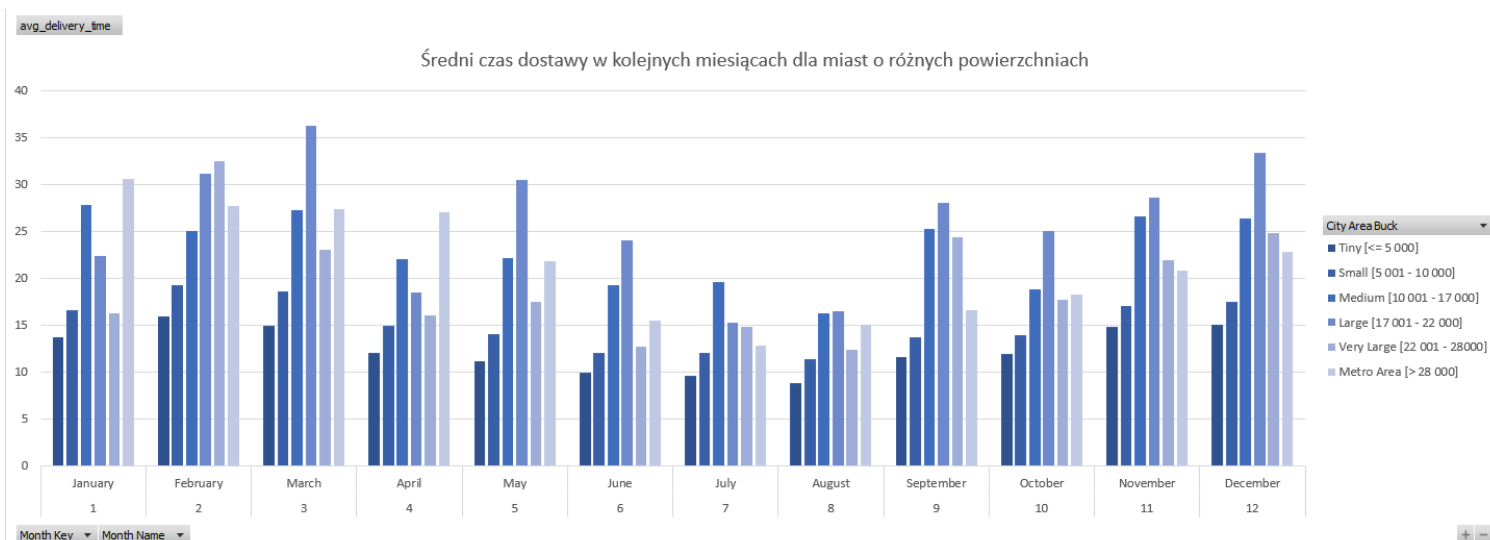
Widać, że największe kwoty padają w regionie południowo-wschodnim, co może być spowodowane większą zamożnością tego regionu jak i większą jego populacją (konsekwencją czego mają więcej różnych zamówień stamtąd i jest większa szansa na jakieś „duże” zamówienie) – rekord płatności padł w trzecim kwartale 2017 roku w regionie południowo wschodnim i wynosił blisko 14000 R\$ (reali brazylijskich) co w przybliżeniu wynosi około 9163,63zł.

#### 4. Średni czas dostawy w kolejnych kwartałach kolejnych lat dla różnych metod płatności



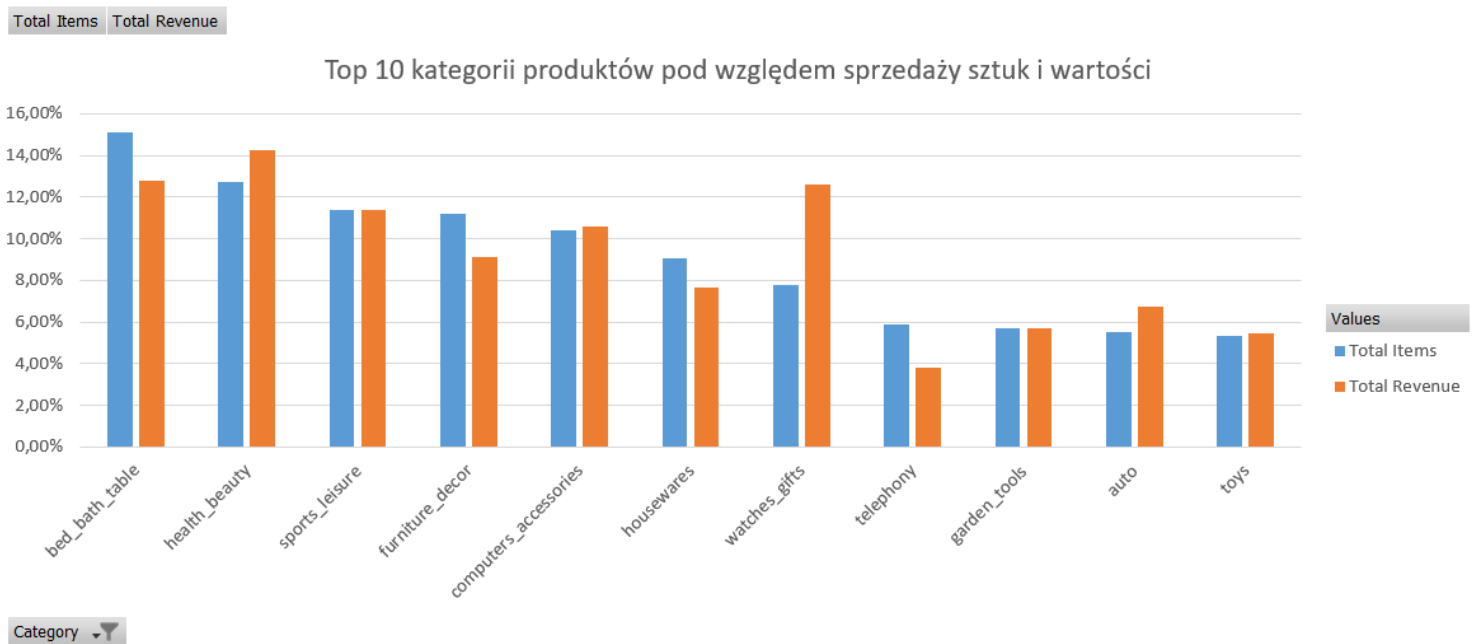
Widać, że czas dostawy spada wraz z biegiem czasu, jest też zbliżony do siebie niezależnie od metody płatności, z lekkim wskazaniem na większe wartości w przypadku boleto (co w wolnym tłumaczeniu znaczy bilet) – możliwe że zaksięgowanie tej formy płatności zajmuje czas, przez co wydłuża się łączna ilość dni od momentu dokonania płatności do otrzymania przesyłki, jednorazowy rekord padł dla vouchera w 2016 roku, ale może to być outlier ponieważ w tym okresie była to jedyna płatność tego typu.

#### 5. Średni czas dostawy w kolejnych miesiącach dla miast o różnych powierzchniach



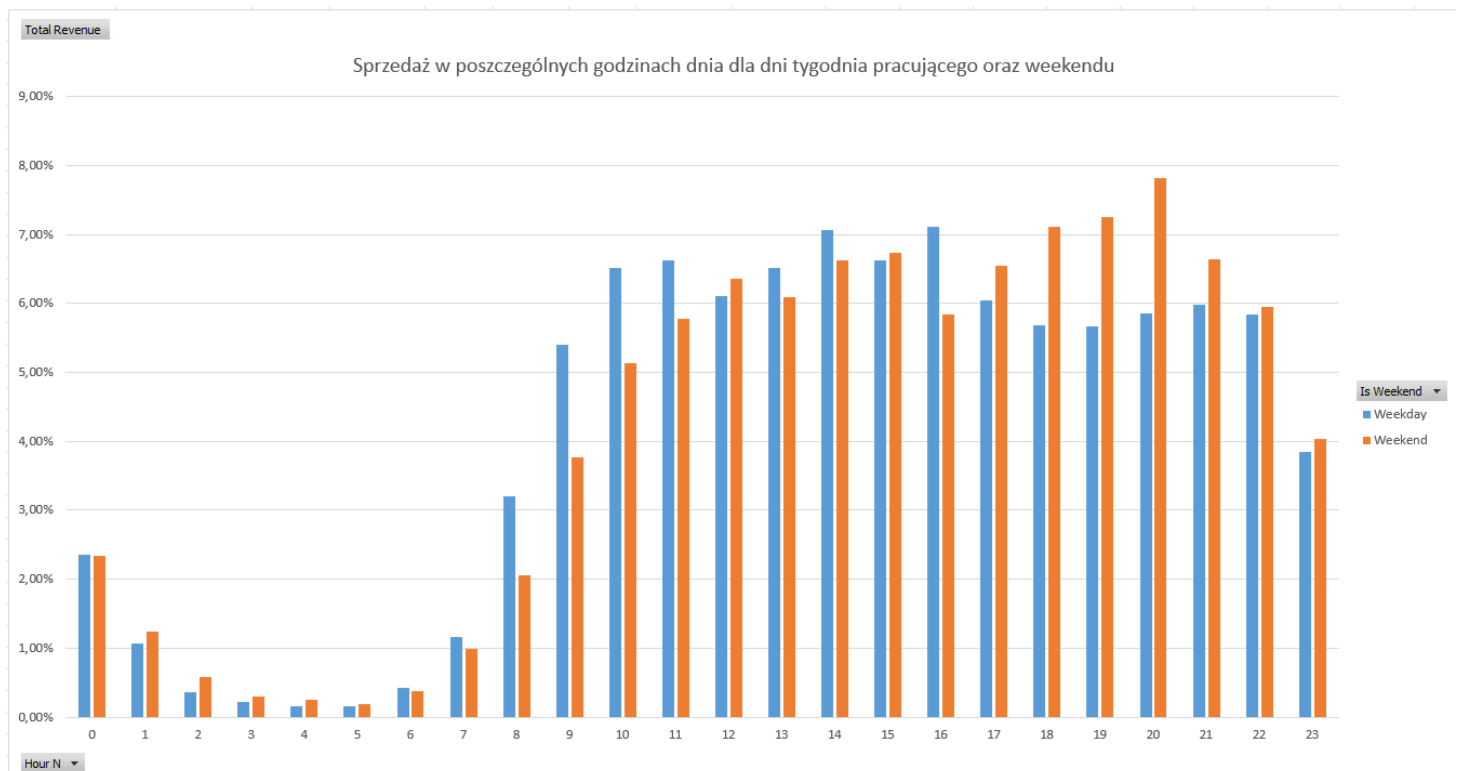
Widać, że dla miast o największej i najmniejszej gęstości zaludnienia czas dostawy jest najszybszy, natomiast dla miast o średniej gęstości zaludnienia widać wzrost w czasie dostawy. Dodatkowo najmniejszy czas dostawy jest w miesiącach letnich, gdzie nawet te największe wartości zazwyczaj mieszczą czas dostawy poniżej dwóch tygodni.

## 6. Top 10 kategorii produktów pod względem sprzedaży sztuk i wartości



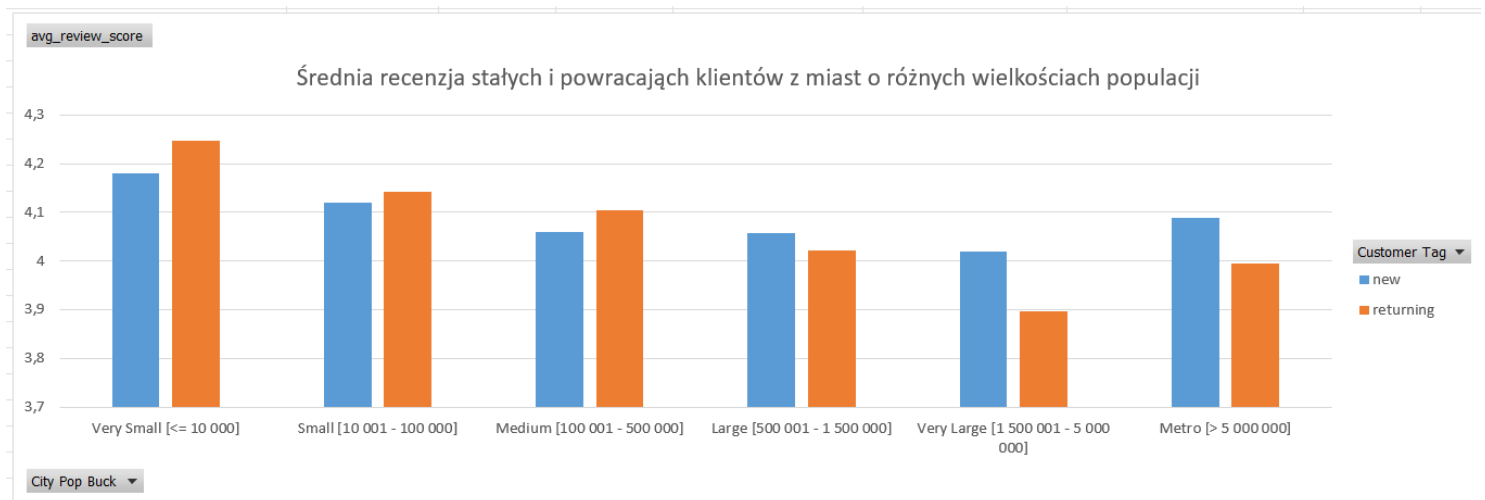
Widać, że sprzedaż ilościowa kategorii często idzie w parze z ich sprzedażą kapitałową, z wyjątkiem watches\_gifts które to kosztują znacznie więcej niż sprzedawane jest ich ilościowo, widać więc że są to przedmioty o większej wartości, które klienci kupują rzadziej.

## 7. Sprzedaż w poszczególnych godzinach dnia dla dni tygodnia pracującego oraz weekendu



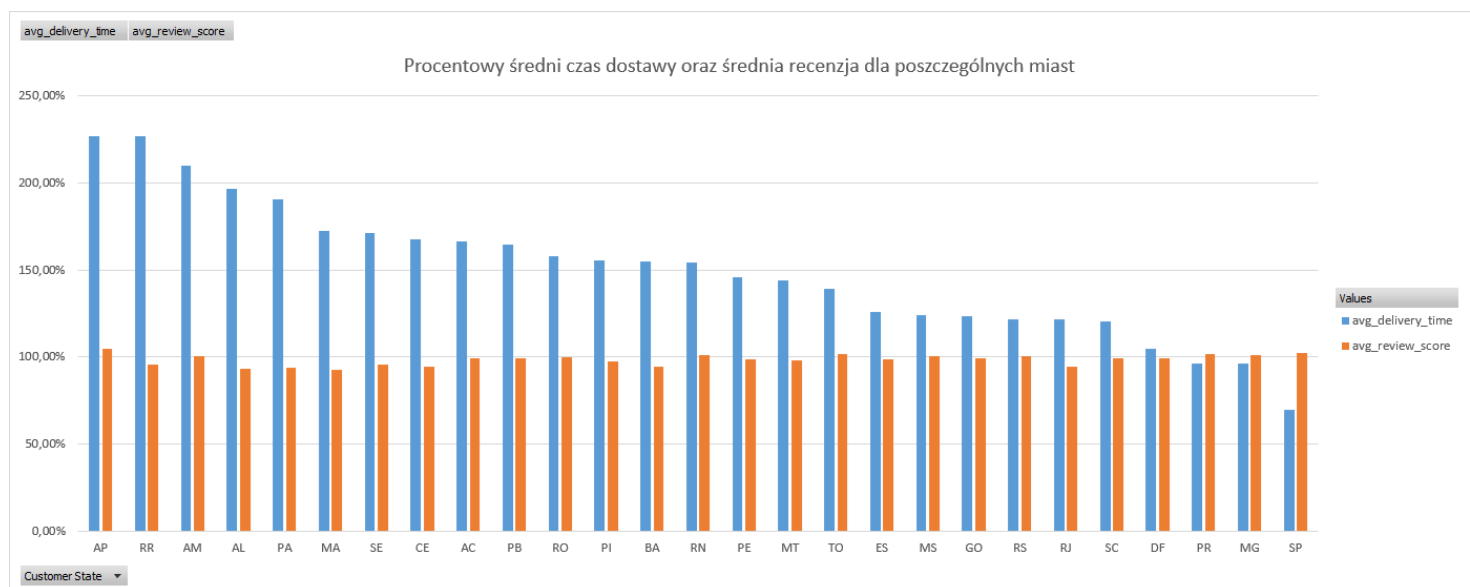
Widać, że godziny szczytu sprzedaży rozpoczynają się około godziny 9:00 rano w przypadku dni od poniedziałku do piątku, natomiast w weekendy widać nieco większy pik w godzinach wieczornych, około godziny 20:00 widać znaczący pik, którego nie ma w tygodniu, może to być spowodowane większą ilością wolnego czasu klientów weekend, przez co decydują się oni robić zamówienia wieczorem.

## 8. Średnia recenzja stałych i powracających klientów z miast o różnych wielkościach populacji



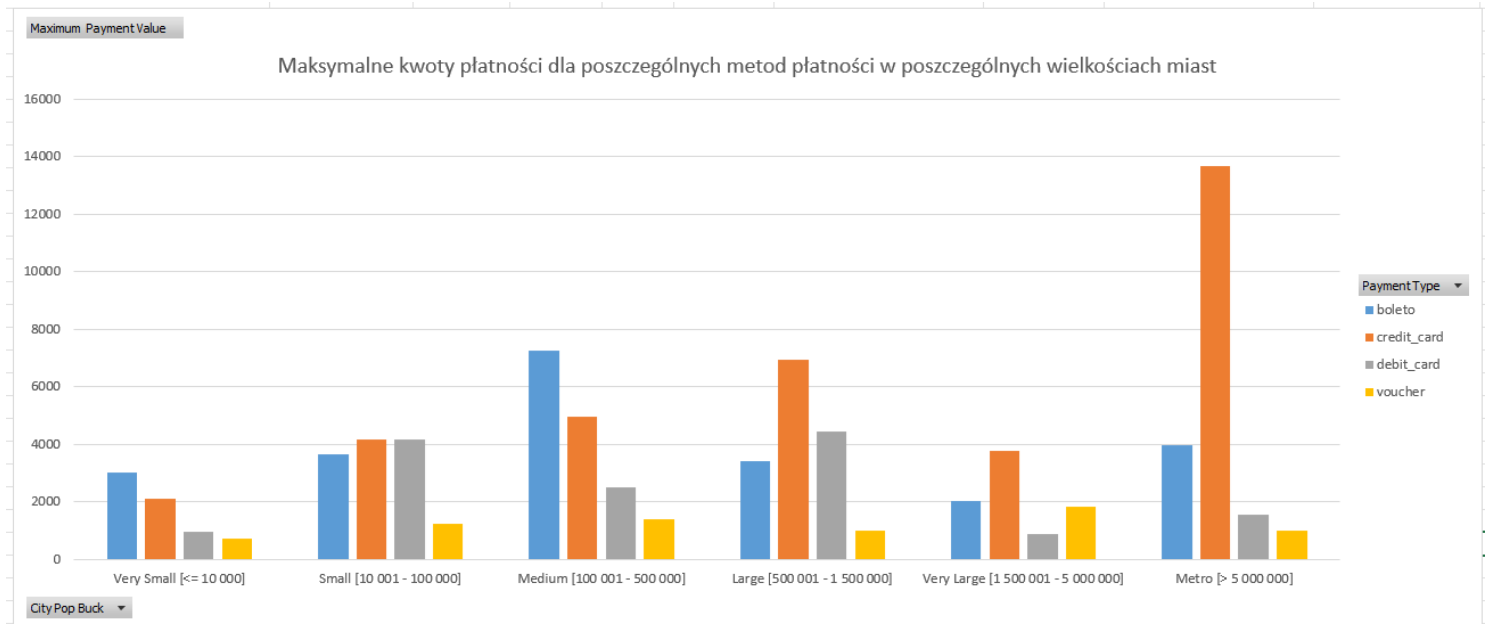
Widać, że nowi klienci zostawiają raczej tym gorsze noty, im w większym mieście mieszkają, natomiast klienci powracający czyli tacy którzy dokonali więcej niż jednej płatności mają na ogół lepsze opinie o sklepie w mniejszych miastach (do 500.000), a w tych większych trend się odwraca i stali klienci oceniają gorzej swoją satysfakcję z zakupów.

## 9. Procentowy średni czas dostawy oraz średnia recenzja dla poszczególnych stanów



Widać, że niektóre stany mają znacznie zawyżony średni czas dostawy (niektóre nawet do blisko 250% średniego czasu dostawy), natomiast inne (raczej te większe takie jak Sao Paulo albo Minas Gerais) mają znacznie szybszą dostawę od reszty (rekordowa dla Sao Paulo, które jest bardzo zaludnionym regionem, więc czas dostawy jest tutaj w zrozumiały sposób mniejszy – natomiast warto zaznaczyć że nie wpływa to na satysfakcję klientów z obsługi i ich ocena pozostaje raczej w okolicach średniej niezależnie od czasu dostawy).

## 10. Maksymalne kwoty płatności dla poszczególnych metod płatności w poszczególnych wielkościach miast



Widać, że dla dużych miast przeważa znacząco karta kredytowa, natomiast w miastach mniejszych i średnich nie jest to już tak oczywiste, a często nawet ludzie płacą największe kwoty alternatywnymi metodami płatności takimi jak bilet czy voucher – warto zaznaczyć tutaj że rozmiary kubelków nie są równomierne, więc dodatkowy czynnik taki jak ilość zamówień z miast o danej populacji może mieć wpływ na te wyniki. Wciąż można jednak wysunąć wniosek, że mieszkańcy mniejszych miejscowości nie robią tak dużych kwotowo zamówień, jak ci z miast o populacji większej niż 5mln.

### Wnioski:

Analizy poprzez powyższe zestawienia pokazują, jak dzięki hurtowni danych i kostce OLAP można szybko identyfikować kluczowe i powtarzalne wzorce sprzedaży (sezonowość, top-kategorie, godziny szczytu), efektywność operacyjną (czasy dostawy w regionach i metodach płatności) oraz zachowania klientów (nowi vs. powracający, preferencje zakupowe wg wielkości miast). Dzięki temu menedżerowie Olist mogą optymalizować logistykę, planować kampanie marketingowe w najlepszych okresach, segmentować klientów i priorytetyzować inwestycje w obszary i kanały przynoszące największe przychody.