

# Hurtownie danych ćwiczenia 1

Ocena końcowa z ćwiczeń – 50% wykonywanie zadań, 50% projekt indywidualny

Zasady wykonywania zadań i ich oceniania:

- Każdy student zakłada repozytorium na Git specjalnie na potrzeby przedmiotu gdzie udostępnia prowadzącemu kod napisany w celu wykonania ćwiczenia i zrzut ekranu efektu końcowego, np. stworzonej tabeli. Termin wykonania i opublikowania ćwiczenia na Git to 23:59 dnia poprzedzającego następne zajęcia. Zadania z każdego ćwiczenia zamieszczone będą przez studentów w osobnych folderach, np. folder do zadań z ćwiczeń numer 1 nazwany będzie „zadania 1”.
- Na następnych zajęciach każde z wykonanych zadań będzie przedstawione przez studentów – jedno zadanie pokazuje jeden student.
- Zadania zostaną sprawdzone przez prowadzącego przed zajęciami. Wykonanie zadań na dane ćwiczenia wiąże się z otrzymaniem plusa. Za nie wykonanie zadań student dostaje minusa.
- Przykładowo – jeżeli student na 9 zajęć raz nie wykonał zadań – jego ocena będzie obliczona jako: 8/9, czyli ~89% co w skali ocen AGH oznacza 4.5.

Projekt indywidualny to projekt który będzie łączył zagadnienia poznane przez studentów podczas wykonywania ćwiczeń, między innymi:

- Wykonanie modelu logicznego i fizycznego hurtowni danych
- Wykonanie ETL wypełniającego tabelę faktów i wymiarów za pomocą danych źródłowych w różnych formatach.
- Wykonaniu zapytań raportowych w języku SQL.
- Wykonaniu raportu PowerBI opartego o dane ze stworzonej wcześniej hurtowni danych.

## Zadanie 1

Na podstawie dostarczonych danych proszę wyznaczyć średnią kwotę zamówienia w dniu 18 Lutego 2015.

## Zadanie 2

Pizzeria planuje wdrożenie programu lojalnościowego dla klientów którzy nigdy nie zamówili pizzy z ananasek w Marcu 2015. Proszę o stworzenie tabeli z id takich zamówień. (przydatna funkcja *string\_agg*)

## Zadanie 3

Pizzeria planuje nagrodzić klientów o najwyższych zamówieniach bonami kwotowymi. Proszę przygotować tabelę z id 10 największych zamówień dla lutego wraz z ich kwotą przy użyciu funkcji *rank () over*.

## Zadanie 4

Proszę stworzyć tabelę która pokaże kwotę każdego zamówienia w jednej kolumnie wraz z id zamówienia w drugiej, oraz średniej kwocie zamówienia dla każdego miesiąca w formacie jak poniżej (w tym przypadku Common Table Expressions mogą okazać się szczególnie przydatne):

	order_id	order_amount	average_month_amount	date
1	2917	40.50	38.022195	2015-02-18
2	5834	33.00	37.402001	2015-04-08
3	11407	45.25	36.706692	2015-07-10
4	14324	66.75	36.430635	2015-08-27
5	19897	71.50	37.767857	2015-12-04
6	1566	20.25	37.112574	2015-01-27
7	7185	20.25	37.743847	2015-05-01
8	10056	12.00	37.674168	2015-06-18
9	15675	28.00	37.939795	2015-09-19
10	18546	10.50	38.534598	2015-11-12
11	46	58.00	37.112574	2015-01-01
12	215	12.00	37.112574	2015-01-04
13	5788	37.50	37.402001	2015-04-07
14	11453	12.50	36.706692	2015-07-10
15	17026	36.75	38.011239	2015-10-15
16	17195	38.20	38.011239	2015-10-18
17	1397	18.50	37.112574	2015-01-24
18	1520	37.50	37.112574	2015-01-26

### Zadanie 5

Proszę przygotować tabelę z listą pokazującą liczbę zamówień dla danej pełnej godziny w dniu 1 Stycznia 2015 tak jak poniżej (proszę zaokrągać do pełnych godzin w dół tj. 11:59 będzie 11):

	order_count	date	hour
1	2	2015-01-01	11
2	7	2015-01-01	12
3	10	2015-01-01	13
4	7	2015-01-01	14
5	7	2015-01-01	15
6	4	2015-01-01	16
7	8	2015-01-01	17
8	8	2015-01-01	18
9	6	2015-01-01	19
10	5	2015-01-01	20
11	2	2015-01-01	21
12	3	2015-01-01	22

### Zadanie 6

Proszę wykonać tabelę z popularnością danych rodzajów pizzy w miesiącu Styczniu 2015. Ma ona pokazywać ilość sprzedanych rodzajów pizz bez rozróżnienia na jej rozmiary. Tabela ma zawierać nazwę każdej pizzy oraz jej kategorię.

### Zadanie 7

Proszę przygotować tabelę która zobrazuje popularność każdego rozmiaru pizzy w miesiącu Lutym oraz Marcu 2015 w formacie zbliżonym do tego poniższego:

	size	count
1	s	2392
2	l	3205
3	m	2481