



**Fundusze  
Europejskie**  
Wiedza Edukacja Rozwój



Politechnika Wrocławska

**Unia Europejska**  
Europejski Fundusz Społeczny



*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

# Hurtownie danych

## Wprowadzenie

**dr inż. Bernadetta Maleszka**

## WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I INNYCH KOMPETENCJI

- Posiadanie wiedzy w zakresie organizacji systemów bazodanowych ze szczególnym uwzględnieniem modelu relacyjnego
- Podstawowa znajomość języka zapytań SQL

## CELE PRZEDMIOTU

- Opanowanie podstawowej wiedzy i umiejętności posługiwania się operatorami grupującymi SQL oraz funkcjami agregującymi i grupującymi SQL
- Opanowanie podstawowej wiedzy i umiejętności dotyczącej charakterystyk przetwarzania zorientowanego na transakcje (OLTP) oraz przetwarzania zorientowanego na analizę (OLAP)
- Opanowanie podstawowej wiedzy oraz umiejętności posługiwania się hurtownią danych
- Opanowanie podstawowej wiedzy i umiejętności dotyczącej integracji, raportowania oraz wizualizacji danych

# PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

Z zakresu wiedzy:

- student ma podstawową wiedzę związaną z zastosowaniem i organizacją hurtowni danych
- student ma podstawową wiedzę związaną z procesem ETL, raportowaniem oraz analizą danych

# PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

Z zakresu umiejętności:

- student potrafi samodzielnie wykorzystywać podstawowe operatory grupujące oraz funkcje agregujące i grupujące SQL
- student potrafi samodzielnie zaprojektować i zaimplementować podstawowy proces ETL
- student potrafi zaprojektować i zaimplementować prostą hurtownię danych i wykorzystać ją do przygotowania prostych raportów, wizualizacji danych i ich analizy

*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

## Wykład

1	Zajęcia organizacyjne. Wprowadzenie do zagadnień HD i BI.	2
2	Operatory grupujące SQL. Funkcje agregujące i grupujące SQL	2
3	Transakcyjne a analityczne potrzeby, procesy i źródła danych	2
4	Wielowymiarowy model danych - warstwa konceptualna	2
5	Wielowymiarowy model danych - warstwa logiczna	4
6	Wielowymiarowy model danych - warstwa fizyczna	2
7	Podstawy procesu ETL	4
8	Architektura hurtowni danych	2
9	Podstawy MDX	4
10	Podstawy raportowania	2
11	Wizualizacja danych. Webowe panele zarządzania	4

## LITERATURA PODSTAWOWA

- Kimball R., Ross M., The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. Wiley Publishing, 2013.
- Kimball R., Caserta J., The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. Wiley Publishing, 2004.
- Inmon W., Building the Data Warehouse, John Wiley & Sons, New York 2005.
- Jensen C.S., Pedersen T.B., Thomsen C., Multidimensional Databases and Data Warehousing, Morgan & Claypool Publishers series Synthesis Lectures On Data Management, 2010.
- Rainardi V., Building a Data Warehouse With Examples in SQL Server, Apress, 2014.
- Harinath S., Pihlgren R., Lee D.G.-Y., Sirmon J., Bruckner R.M., Professional Microsoft® SQL Server® 2016 Analysis Services With MDX and DAX, John Wiley & Sons, Inc., 2016.

## LITERATURA UZUPEŁNIAJĄCA

- Bhatia P., Data Mining and Data Warehousing. Principles and Practical Techniques. Cambridge University Press, 2019.
- Imhoff C., Galemme N., Geiger J. G., Mastering Data Warehouse Design, Wiley Publishing, Inc., 2003.
- Dela J., Implementing Business Intelligence with SQL Server 2019. Packt Publishing, 2019.



„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”

## Semestr letni

	LUTY	MARZEC					KWIECIEŃ				MAJ				CZERWIEC					LIPIEC	
PN	24	3	10	17	24	31	7	14	21	28	5	12	19	26	2	9 Cz/P	16	23	30	7	14
WT	25	4	11	18	25	1	8	15	22	29	6	13	20	27	3	10	17	24 Pt/P	1	8	15
ŚR	26	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18 Pt/N	25 Cz/N	2	9	
CZ	27	6	13	20	27	3	10	17	24	1	8	15	22 Cz/P	29	5	12	19	26	3	10	
PT	28	7	14	21	28	4	11	18	25	2	9	16	23 Pt/P	30	6	13	20	27	4	11	
SO	1	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	5	12	
N	2	9	16	23	30	6	13	20	27	4	11	18	25	1	8	15	22	29	6	13	
P - PARZYSTY N - NIEPARZYSTY	N	P	N	P	N	P	N	P	N	P	N	P	N	P	N	P	N	P	N	P	N

■ Święta  
 ■ Dodatkowe dni wolne  
 ■ Zmiany  
 ■ Sesja  
 ■ Przerwa międzysemestralna



Fundusze  
Europejskie  
Wiedza Edukacja Rozwój



Politechnika Wroclawska

Unia Europejska  
Europejski Fundusz Społeczny



*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

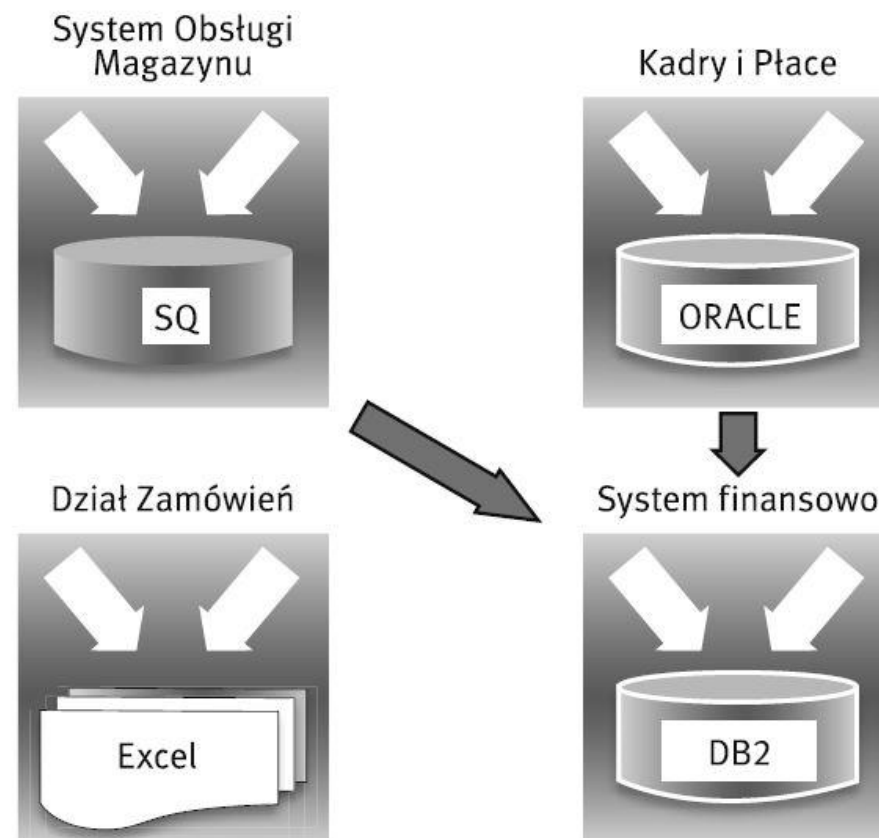
# Zasady zaliczenia

- **Egzamin:**

- I termin: 27.06.2024 r., godz. 7.30 - 9.00, s. 205 C1
- II termin: 04.07.2024 r., godz. 7.30 - 9.00, s. 29 D1

## Historia baz danych

- Różne systemy, różne aspekty działania:
- wystawienia faktur,
- obsługa magazynu,
- systemy kadrowe,
- systemy księgowe,
- obsługa klientów
- ...



# OLTP - Online Transaction Processing

- Przechowywane dane są zorientowane procesowo (np. wystawianie faktury)
- Stosunkowo niewielkie rozmiary baz danych (kilka gigabajtów)
- Przechowywane są dane bieżące bez konieczności gromadzenia danych historycznych
- Realizowana jest duża ilość prostych zapytań
- Przechowywane są dane elementarne
- Realizowane są operacje wstawiania, modyfikowania i usuwania danych



Pozyskiwanie informacji  
potrzebnych kierownictwu  
do podejmowania decyzji?

# OLTP - Online Transaction Processing

- Przechowywane dane są zorientowane procesowo (np. wystawianie faktury)
- Stosunkowo niewielkie rozmiary baz danych (kilka gigabajtów)
- Przechowywane są dane bieżące bez konieczności gromadzenia danych historycznych
- Realizowana jest duża ilość prostych zapytań
- Przechowywane są dane elementarne
- Realizowane są operacje wstawiania, modyfikowania i usuwania danych

NIE WYSTARCZA!!!



## OLTP

### Online Transaction Processing

- Przechowywane dane są zorientowane procesowo (np. wystawianie faktury)
- Stosunkowo niewielkie rozmiary baz danych (kilka gigabajtów)
- Przechowywane są dane bieżące bez konieczności gromadzenia danych historycznych
- Realizowana jest duża ilość prostych zapytań
- Przechowywane są dane elementarne
- Realizowane są operacje wstawiania, modyfikowania i usuwania danych

## OLAP

### Online Analytical Processing

- Przechowywane dane są zorientowane tematyczne (np. sprzedaż produktów, stany zapasów)
- Bardzo duże ilości gromadzonych danych (terabajty, petabajty)
- Przechowywane są dane bieżące i historyczne
- Bardzo złożone zapytania operujące na wielkich ilościach danych
- Przechowywane są dane elementarne i zagregowane (sumy, średnie)
- Wykonywane są głównie operacje dopisywania nowych danych – praktycznie nie wykonuje się operacji modyfikowania danych



Fundusze  
Europejskie  
Wiedza Edukacja Rozwój



Politechnika Wrocławska

Unia Europejska  
Europejski Fundusz Społeczny



*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

# Hurtownia danych - definicja

**Hurtownia danych to:**

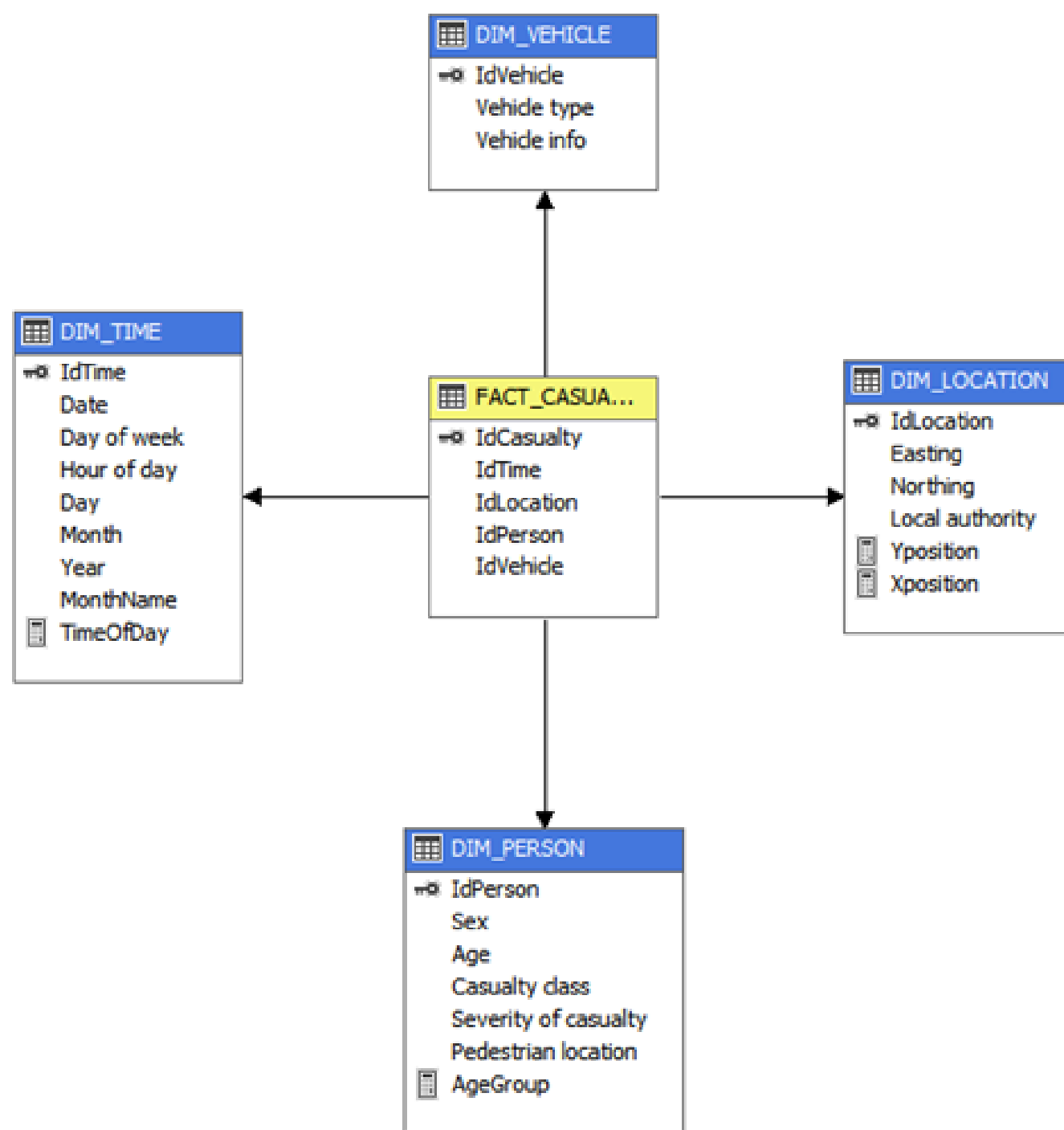
- tematycznie zorientowana
- zintegrowana
- chronologiczna
- trwała

kolekcja danych do wspomagania procesów podejmowania decyzji

## Przykład hurtowni

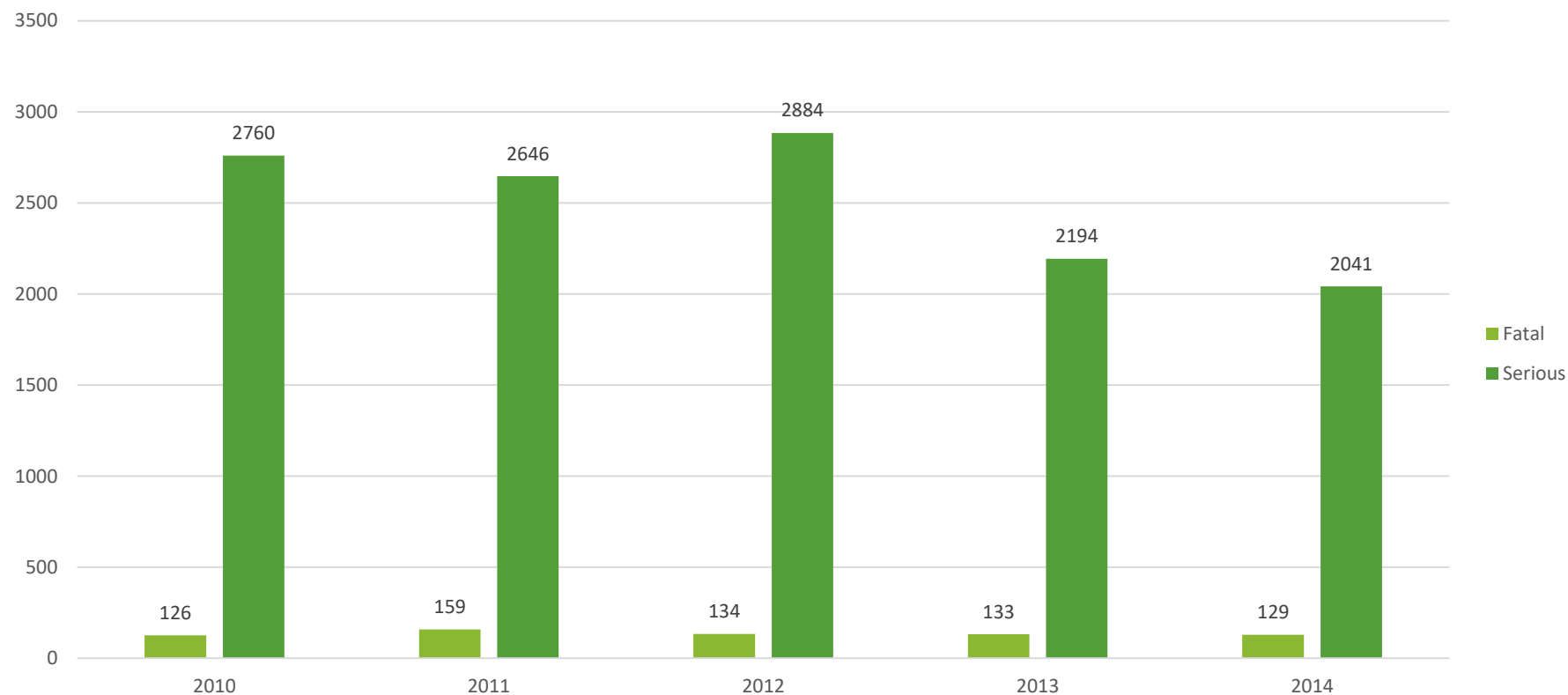
- Dane dotyczące wypadków drogowych
- Cel: analiza **liczby** ofiar wypadków drogowych w zależności od:
  - czasu (daty i pory dnia),
  - danych ofiary,
  - pojazdu sprawcy
  - oraz lokalizacji





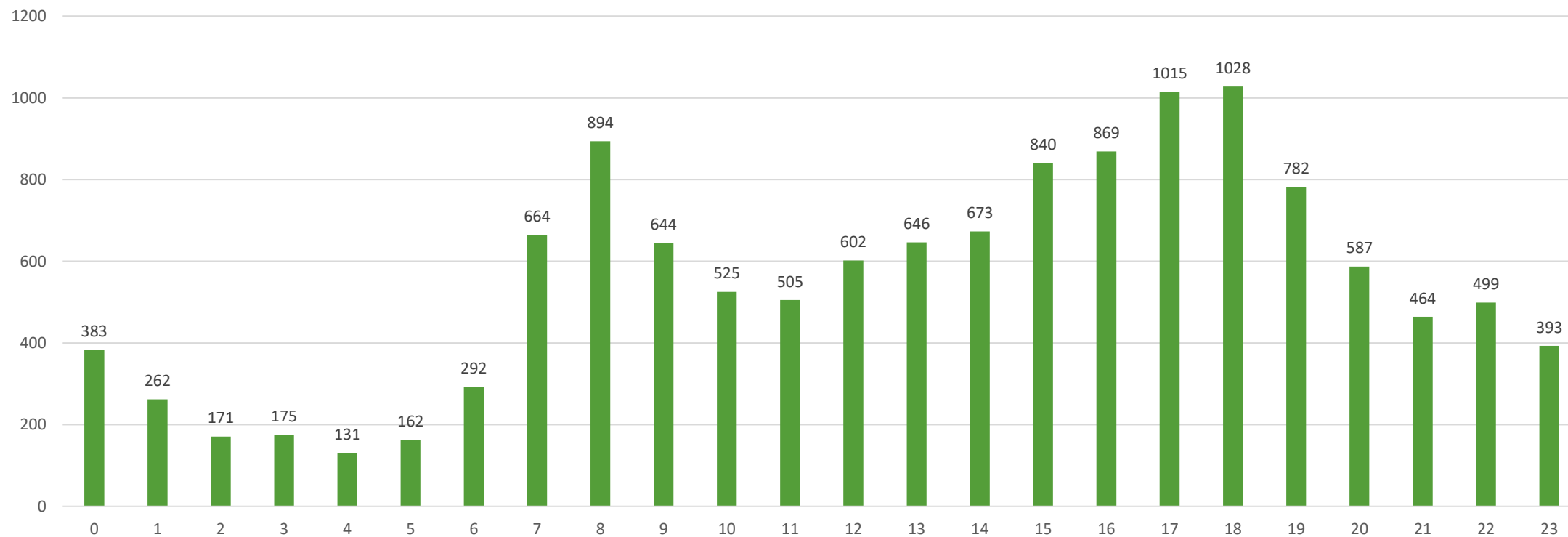
*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

## Liczba ofiar na przestrzeni lat

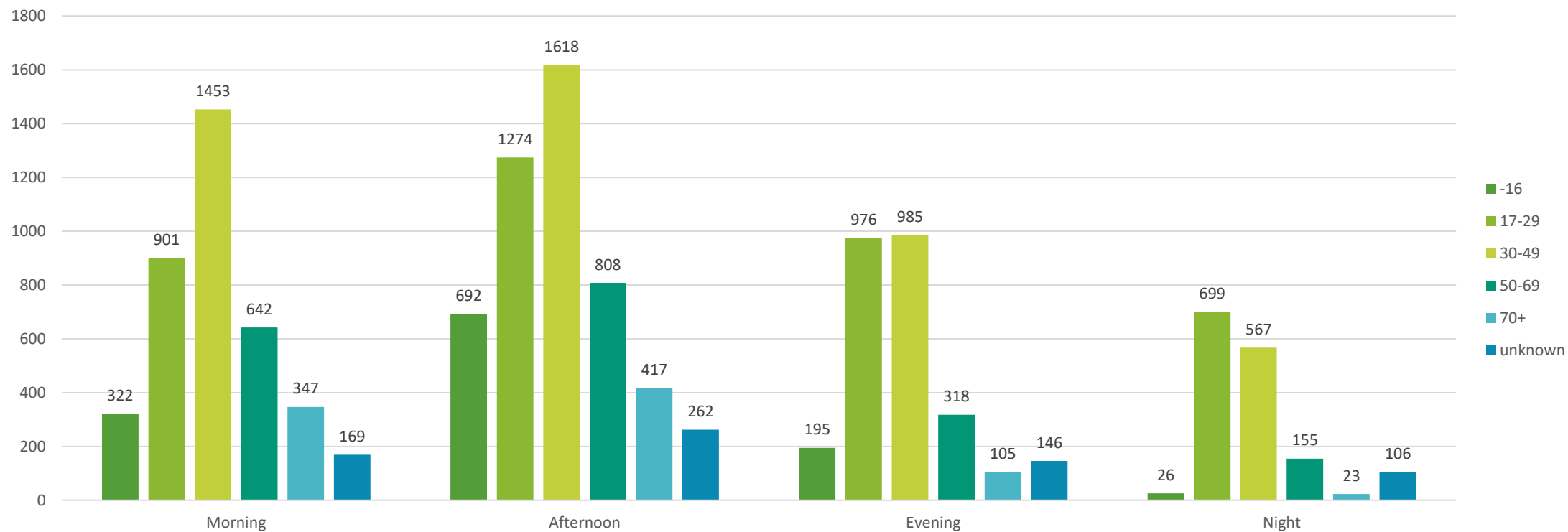


*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

## Liczba ofiar – rozkład w ciągu doby

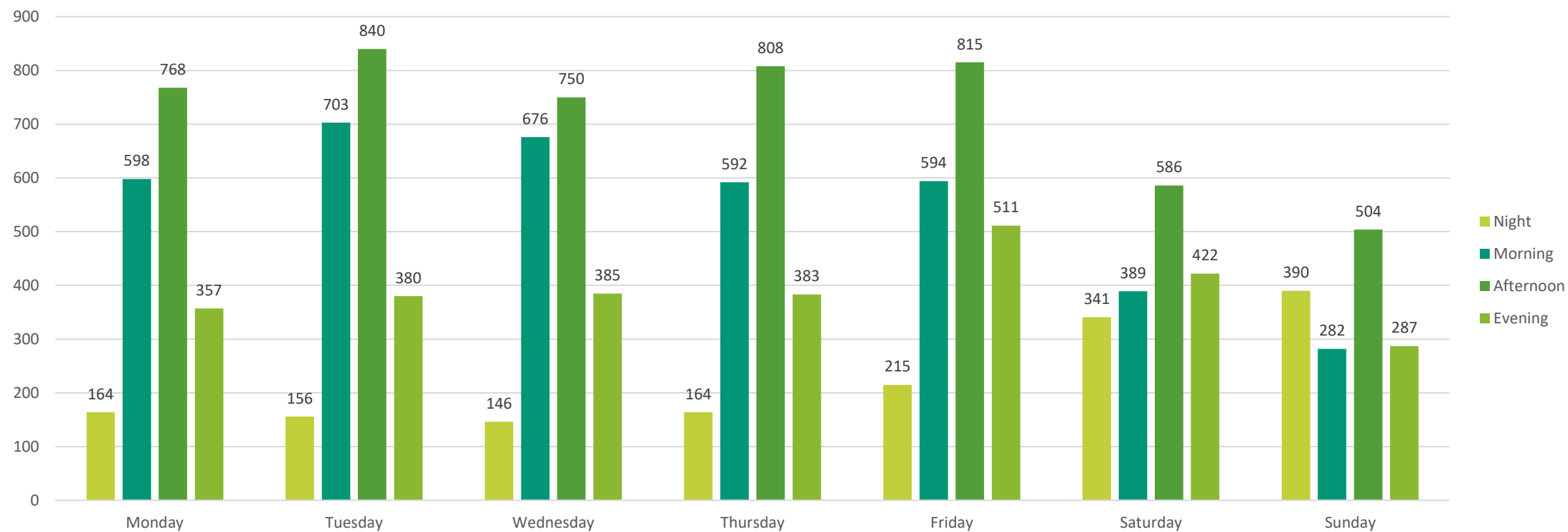


## Liczba ofiar w podziale na wiek ofiary i porę dnia

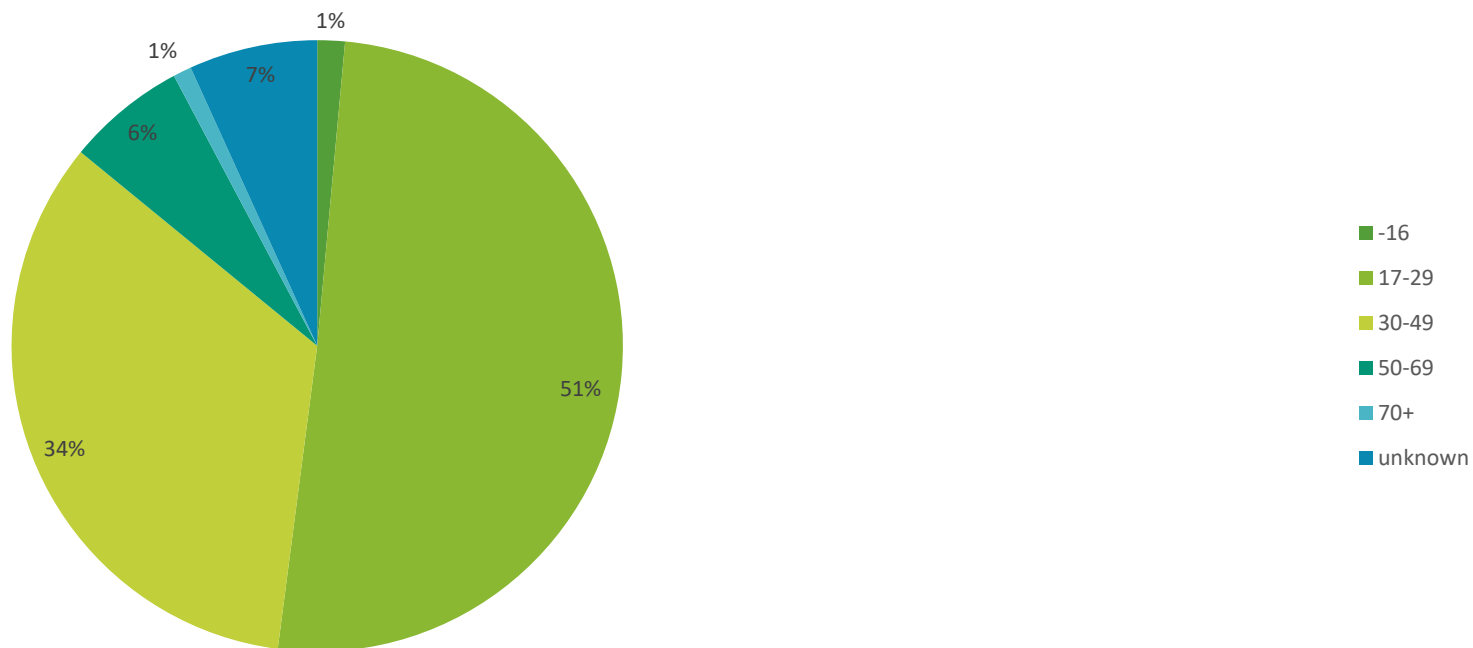


*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

# Liczba ofiar w zależności od pory dnia w ciągu tygodnia

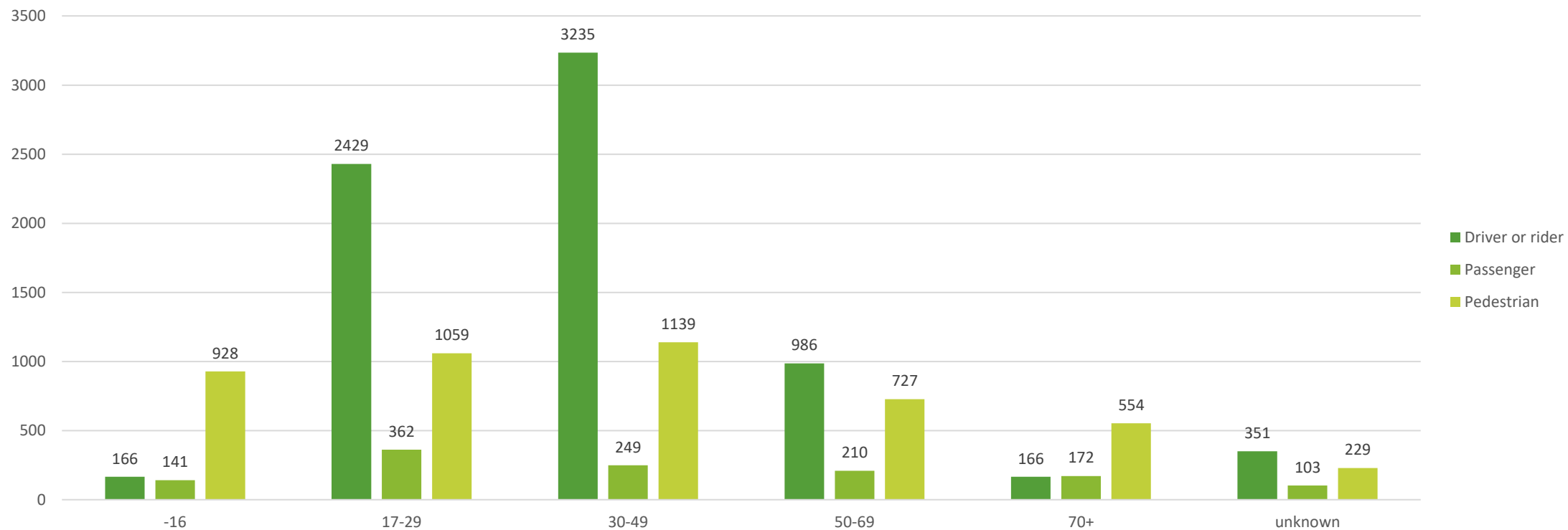


## Procent ofiar w zależności od wieku w weekendowe noce

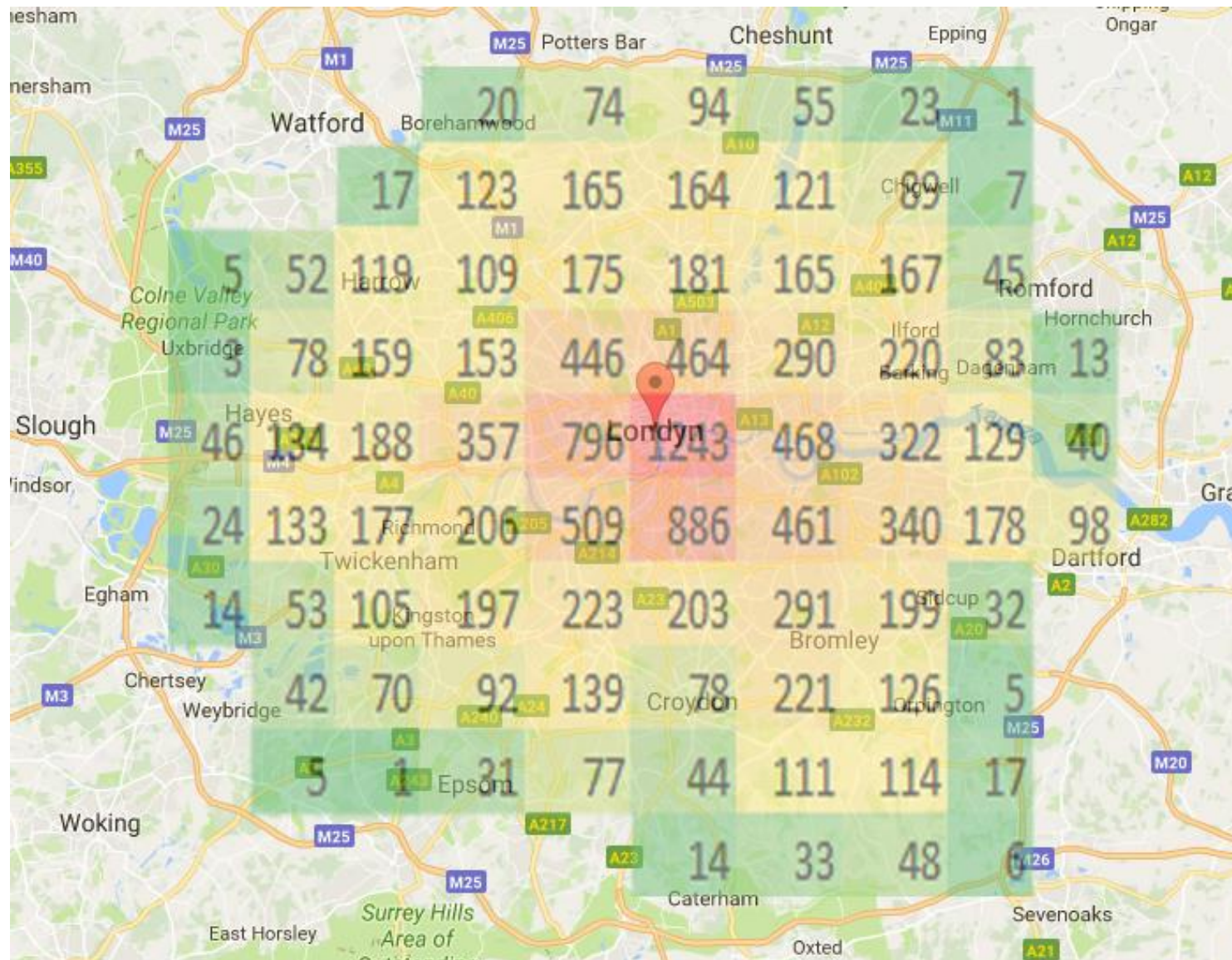


*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

## Rodzaj ofiary w zależności od wieku



**„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”**





## Przykład hurtowni

- Sklep rowerowy
- Transakcje z okresu 3 lat
- **Cel: analiza sprzedaży**, np. *liczby sprzedanych produktów oraz wartości transakcji oraz wartości obniżek/promocji/itp. w zależności od produktu, klienta, sprzedawcy oraz daty transakcji*
- **Ćwiczenie 1:**
- Zaproponuj co najmniej dwuwymiarowe zestawienia (raporty), które będą wspomagały podejmowanie decyzji biznesowych w sklepie rowerowym

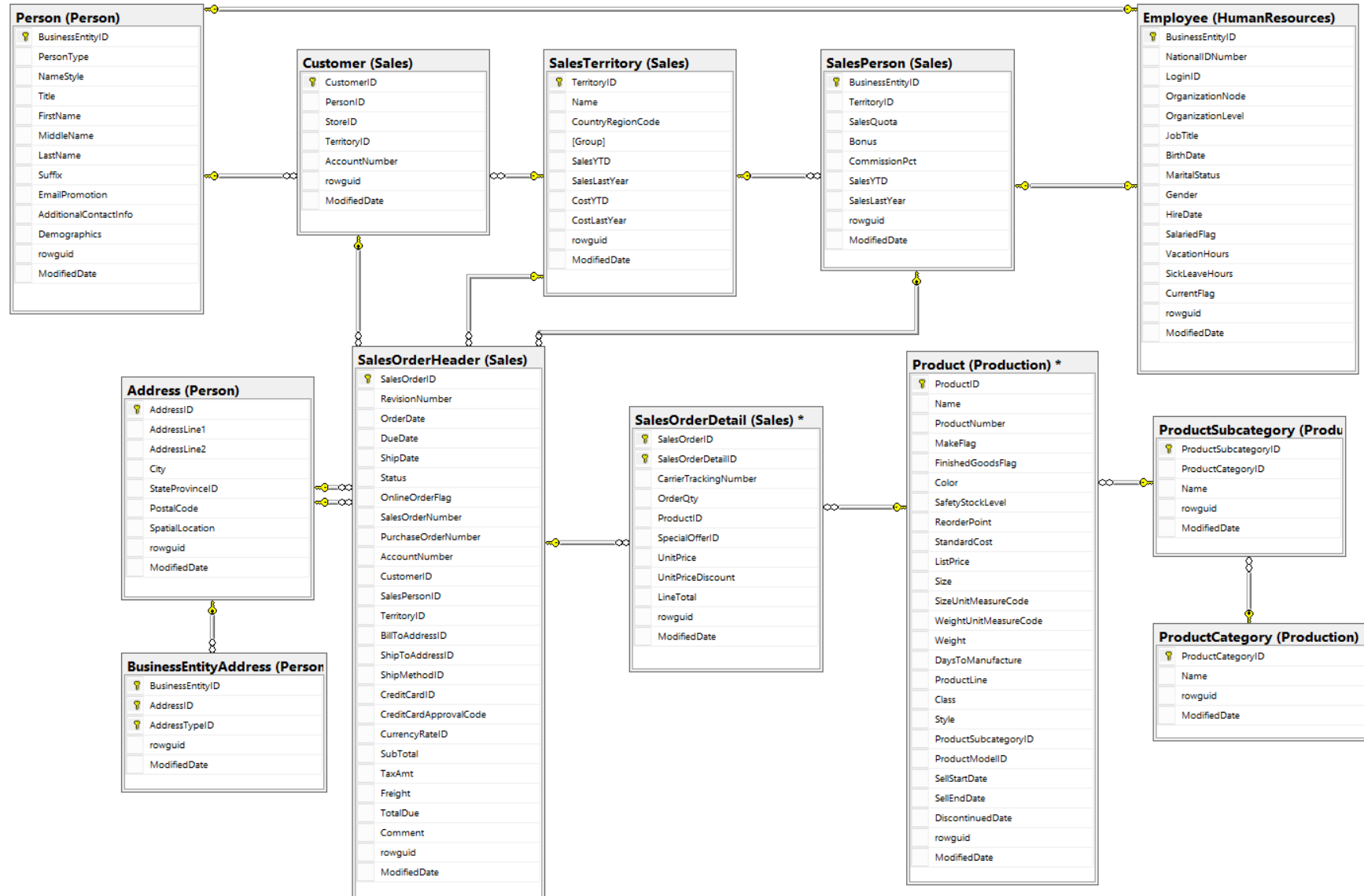


Fundusze Europejskie  
Wiedza Edukacja Rozwój



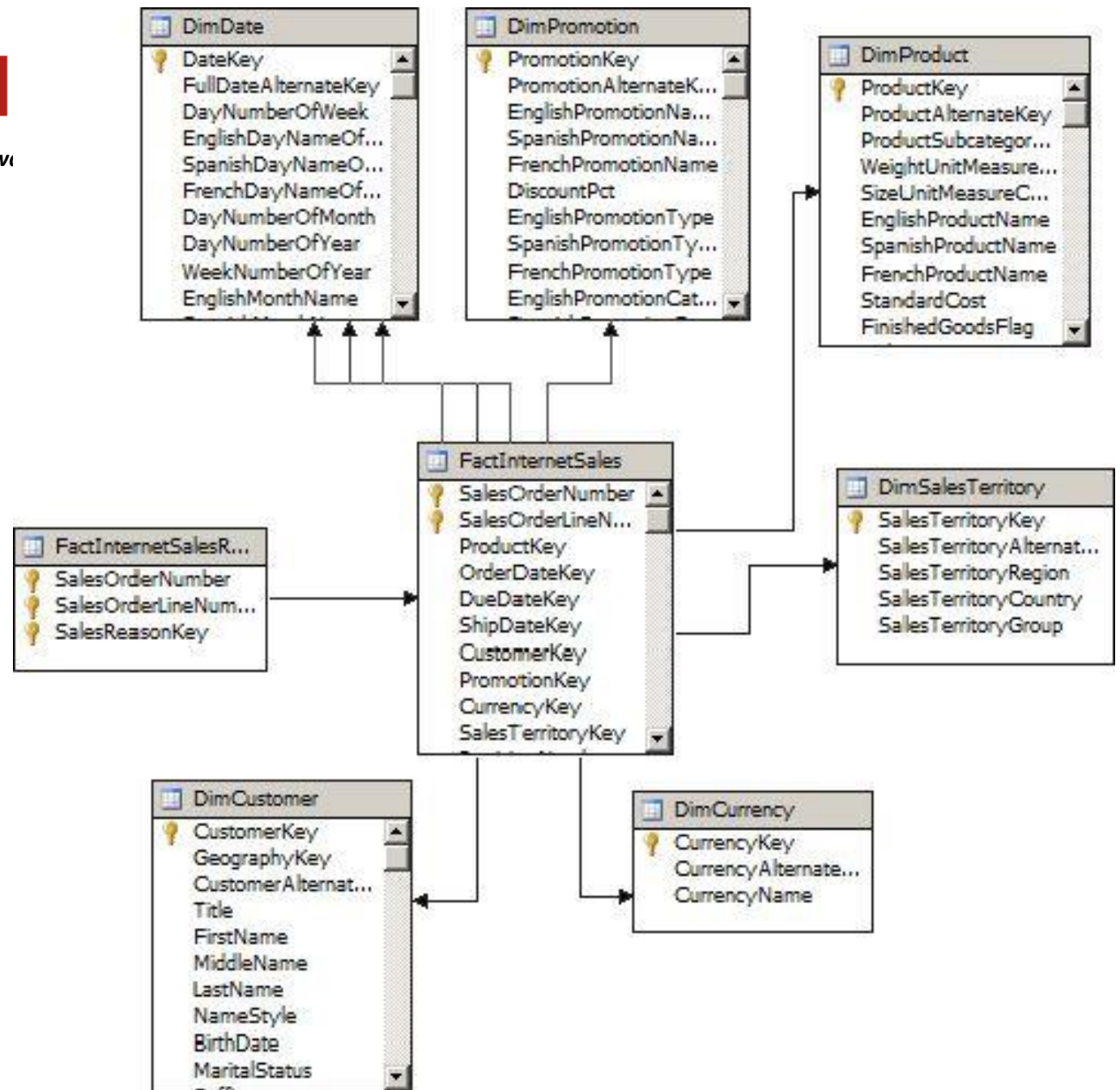
Politechnika Wrocławska

Unia Europejska  
Europejski Fundusz Społeczny



# Zadanie

- Na podstawie przedstawionych tabel z AdventureWorksDW określ:
- Jakie cele biznesowe można zrealizować na podstawie danych zebranych w hurtowni o podanym schemacie?
  - Jakie analizy można przeprowadzić? Zdefiniuj min. 10 co najmniej dwuwymiarowych zestawień, które będą wspomagały podejmowanie decyzji biznesowych





**Fundusze  
Europejskie**  
Wiedza Edukacja Rozwój



Politechnika Wrocławska

**Unia Europejska**  
Europejski Fundusz Społeczny



*„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”*

# Hurtownie danych

**Dziękuję za uwagę**

dr inż. Bernadetta Maleszka