

# Grafika Informacyjna

## Specyfikacja projektu

### *Analiza danych dotyczących wypadków podczas pracy*

Adam Brych 117325

Marcin Halczuk 117275

#### 1. Cel projektu

Celem projektu jest analiza danych dotyczących wypadków podczas pracy oraz przedstawienie ich w przystępniejszej niż obecnie formie poprzez zastosowanie wykresów, kartogramów oraz innych grafik (obraz ciała człowieka). Dzięki temu późniejsza praca nad dostarczonymi danymi, wysuwanie wniosków na ich podstawie, zauważanie ciekawych zależności oraz analiza poszczególnych czynników będzie łatwiejsze oraz mniej czasochłonne.

#### 2. Lista zadań

- 2.1. Wstępne przetworzenie danych do formatu, który umożliwi ich sprawne wczytanie oraz tworzenie wykresów, kartogramów i innych grafik (obraz ciała człowieka).
- 2.2. Stworzenie wykresów słupkowych, które pozwolą na łatwiejszą analizę wartości, proporcji między nimi, znajdowanie wartości maksymalnych i minimalnych.
- 2.3. Wygenerowanie map z podziałem na województwa, na które nanoszone będą wartości takie jak liczba wypadków, procent wypadków śmiertelnych, które pozwolą na łatwiejszą analizę pod kątem regionu.
- 2.4. Stworzenie wykresów słupkowych, które ułatwią porównywanie sektorów gospodarczych pod kątem poszczególnych czynników.
- 2.5. Przedstawienie danych odnośnie części ciała, które uległy wypadkowi na mapach cieplnych.
- 2.6. Wygenerowanie wykresów, na które zostanie naniesionych kilka czynników, co ułatwi ich analizę w danym sektorze.
- 2.7. Stworzenie wykresów zawierających linię trendu, które ułatwią odnajdywanie ciekawych zależności, takich jak liczba wypadków w zależności od wieku oraz doświadczenia pracownika.
- 2.8. Utworzenie wykresów pokazujących zmianę poszczególnego czynnika na przestrzeni lat dla danych sektorów, w celu usprawnienia szukania zależności w czasie.

### 3. Źródło danych i ich charakterystyka

Analizowane dane zamieszczone są na stronie Głównego Urzędu Statystycznego w obszarze tematycznym Rynek pracy – Warunki pracy<sup>[1]</sup>. Dostępne jest archiwum z 10 ostatnich lat (2016-2007) w postaci plików XLS, osobnych dla każdego roku. Wewnątrz pliku znajdują się arkusze kalkulacyjne zawierające dane w postaci tabelarycznej. Każda tabela dotyczy innego czynnika oraz jest podzielona na sektory gospodarcze, każdy czynnik zamieszczony jest na osobnym arkuszu. Wewnątrz tabeli podane są liczby wypadków dla poszczególnego sektora oraz czynnika, dodatkowo każdy wypadek jest przypisany do jednej z trzech kategorii (lekki, ciężki i śmiertelny) np. liczba wypadków podzielonych na lekkie, ciężkie oraz śmiertelne dla poszczególnych sektorów takich jak górnictwo, rolnictwo, budownictwo. Dla każdego czynnika podano także sumę wydarzeń po wszystkich sektorach.

### 4. Wybór czynników

- a) Podział wypadków ze względu na miesiące (lata) oraz sektory gospodarcze i województwa. Zdecydowano się wybrać ten czynnik ze względu na możliwość pokazania ciekawych zależności liczby wypadków w zależności od pory roku.
- b) Podział wypadków ze względu na staż pracy oraz sektor gospodarczy. Czynnik ten może dobrze zobrazować na przestrzeni czasu poziom świadomości zagrożeń podczas wykonywania zawodu.
- c) Podział typu wypadku (lekki, ciężki, śmiertelny) na przestrzeni 10 lat dla wybranych sektorów – ukazanie postępu technologicznego, który pozwolił na zmniejszenie liczby wypadków.
- d) Wypadki w zależności od wieku pracownika, płci oraz sektora gospodarczego – pomoc we wskazaniu obszarów środowiska pracy, które wymagają dostosowania ergonomicznego do warunków fizycznych pracownika.
- e) Umieszczenie urazu – trudność w analizie danych w postaci tabelarycznych.
- f) Suma wypłaconych odszkodowań.

### 5. Zarys projektu

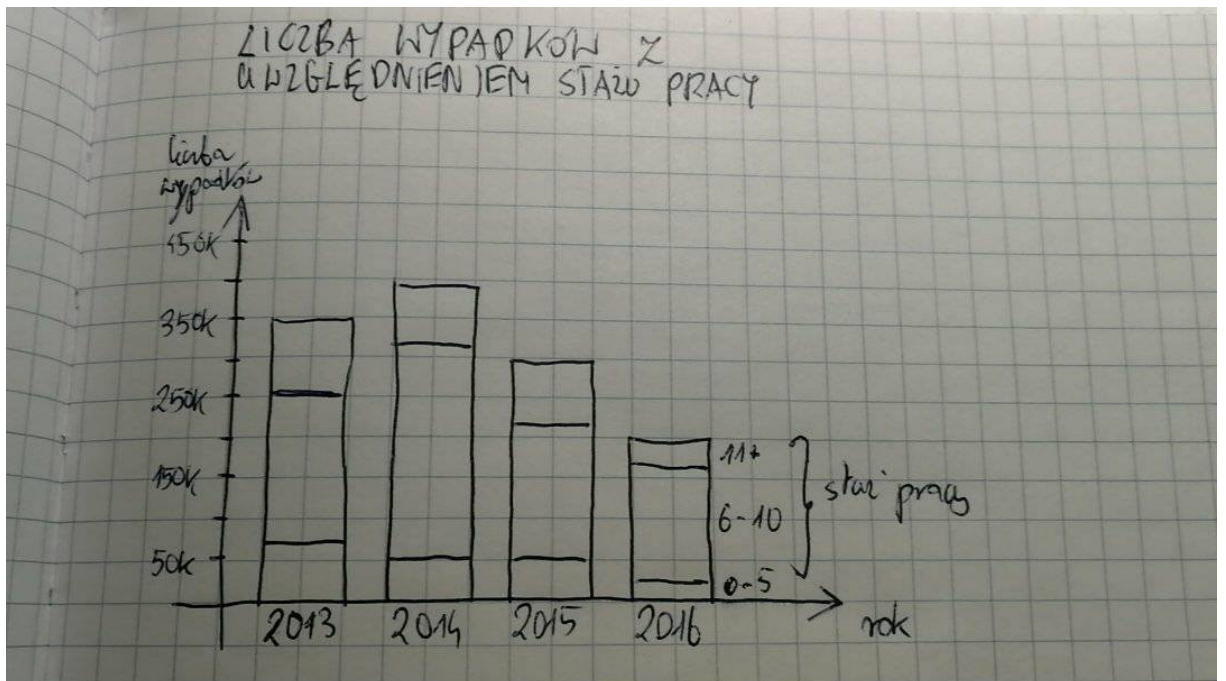
#### 5.1. Środowisko programistyczne

Zdecydowano się na wykorzystanie języka programowania Python wykorzystującego biblioteki Gnuplot<sup>[2]</sup> oraz Matplotlib<sup>[3]</sup> ze względu na oferowany przez nich szeroki wachlarz wykresów i możliwości ich dostosowania. Dodatkowo język Python wydaje się dobrym wyborem biorąc pod uwagę przetwarzanie dużej liczby danych tabelarycznych.

## 5.2. Aplikacja

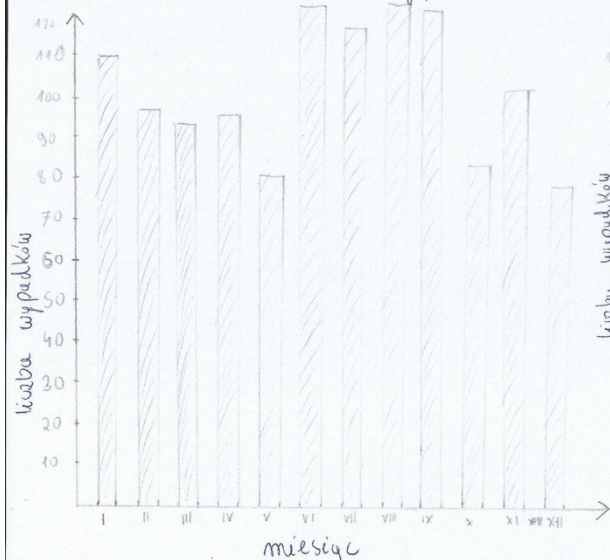
W ramach projektu stworzona zostanie aplikacja konsolowa, która jako wejście przyjmować będzie pliki XLS (jeden lub więcej) pobrane bezpośrednio z Głównego Urzędu Statystycznego. Na wyjściu program będzie generować raport w postaci pliku PDF zawierający wykresy, kartogramy oraz grafiki omówione w punkcie 2. stworzone na podstawie podanych na wejściu danych. Dodatkowo możliwe będzie zawężenie raportu do wykresów dotyczących wybranych czynników (zbiór z góry określony przez twórców aplikacji, wymienione w punkcie 4.) poprzez podanie przełączników.

## 5.3. Rys

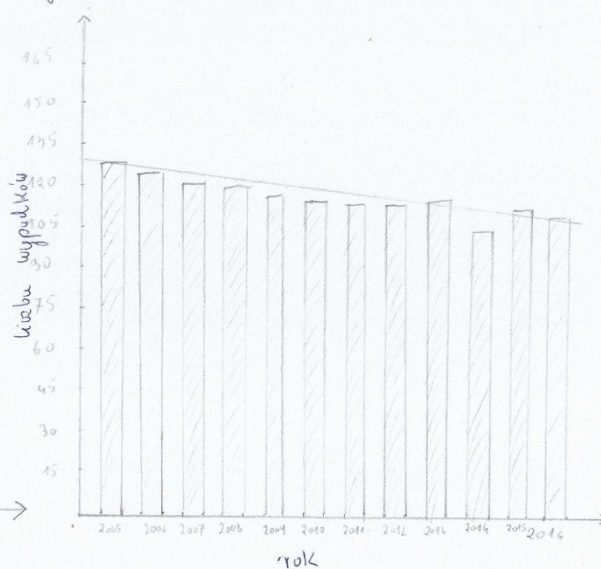


unki koncepcyjne

### Analiza wypadków ze względu na miesiąc

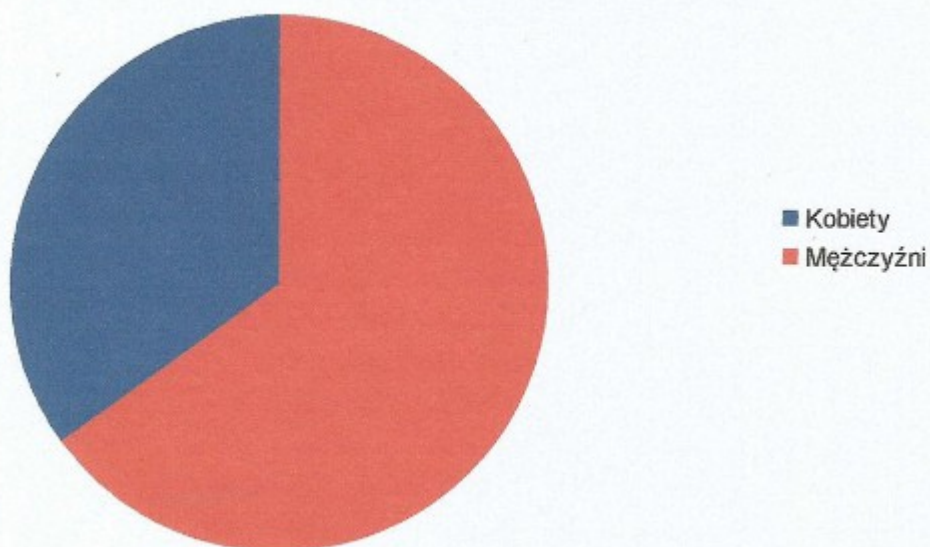


Wykres wypadków w rolnictwie, leśnictwie, łowiectwie i rybactwie w 2016 roku z podziałem na miesiące



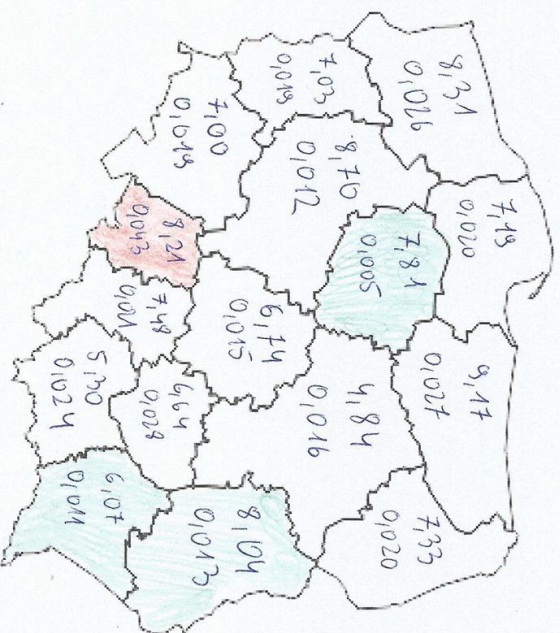
Wykres wypadków w rolnictwie, leśnictwie, łowiectwie i rybactwie w miesiącach styczeń na październik lat 2005-2016

### Procentowy udział wypadków z podziałem na płeć



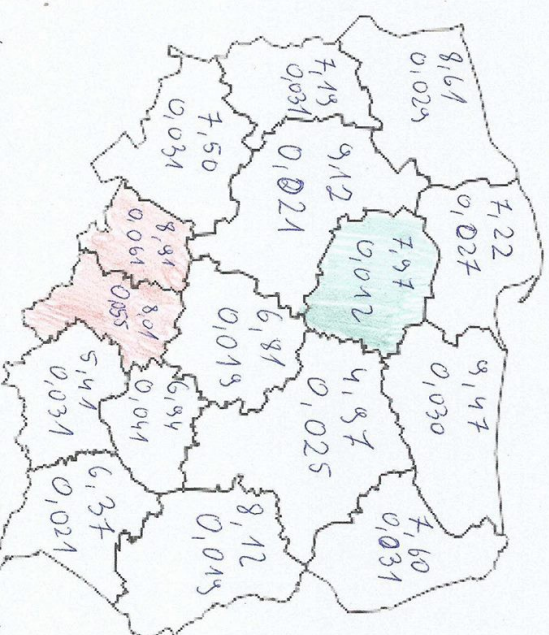


# Poskolowani w wypadkach przy pracy według województwa



liczba wypadków ogółem i śmiertelnych  
w 2016 roku

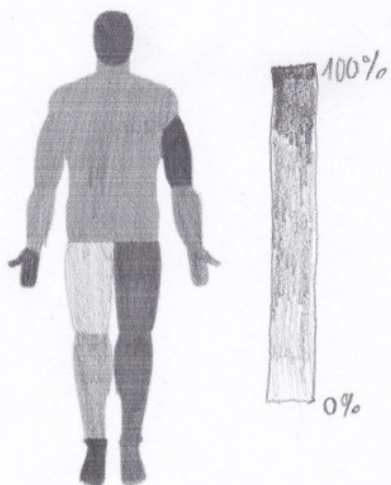
- w podzieleniu na 1000 pracujących
- - liczba wypadków śmiertelnych powyżej 0,015
- - liczba wypadków śmiertelnych powyżej 0,040



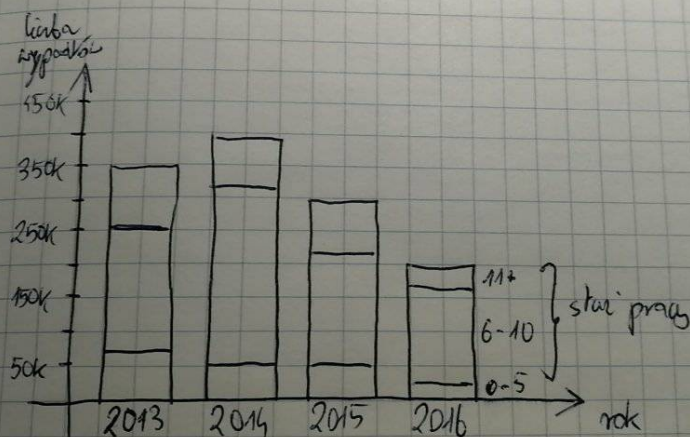
średnia liczba wypadków ogółem i śmiertelnych  
2 lat 2005 - 2016 w podzieleniu  
na 1000 pracujących

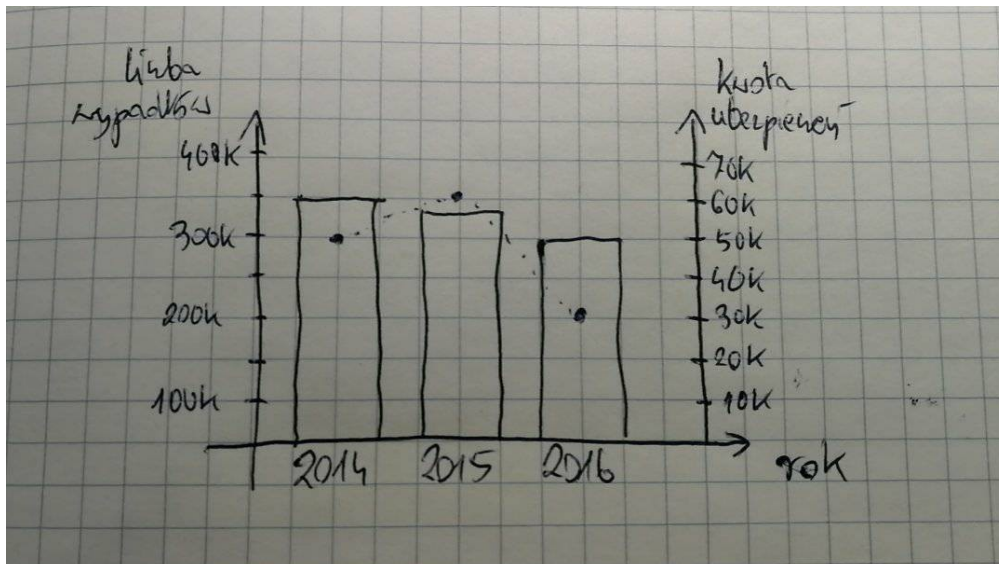
- - liczba wypadków śmiertelnych powyżej 0,015
- - liczba wypadków śmiertelnych powyżej 0,040

# UMIEJSCOWIENIE URAZU



## LICZBA WYPADKÓW Z UZGLĘDNIENIEM STAŻU PRACY





## 6. Harmonogram działań

27.11. – przekonwertowanie i wczytanie danych z plików xls do przygotowanych struktur danych w języku Python.

04.11. – wygenerowanie wykresów słupkowych dotyczących czynników na przestrzeni lat.

18.12. – utworzenie map cieplnych dla danych zależnych od trzech zmiennych oraz danych o umiejscowieniu urazów.

08.01. – przygotowanie grafik w postaci mapy Polski wraz z województwami i naniesionymi na nie wartościami czynników.

Styczeń 2018 – przygotowanie dokładnego opisu analizowanych danych wraz z wnioskami.

1. Główny Urząd Statystyczny, Wypadki przy pracy. <http://stat.gov.pl/obszary-tematyczne/rynek-pracy/warunki-pracy-wypadki-przy-pracy/>
2. Michael Haggerty, Gnuplot.py. <http://gnuplot-py.sourceforge.net/>
3. John Hunter, Darren Dale, Eric Firing, Matplotlib. <https://matplotlib.org/>