

Data oddania: _____

Ocena: _____

Mateusz Walczak 216911

Konrad Kajszczak 216790

Zadanie 2: Lingwistyczne podsumowania baz danych*

1. Cel

Praca w toku

2. Wprowadzenie

Praca w toku

3. Opis implementacji

Praca w toku

4. Materiały i metody

Wybrana przez nas baza danych zawiera historyczne pomiary pogodowe z Holandii [1]. Dane zostały zgromadzone przez KNMI (*Dutch weather institute* - Holenderski instytut pogody) na przestrzeni lat 1901-2018 i pochodziły z 50 różnych stacji pogowych znajdujących się na terenie całego kraju.

* SVN: <https://github.com/Walducha1908/KSR2>

Ze względu na fakt, iż oryginalna baza danych składa się z 804099 krotek, postanowiliśmy wybrać tylko niewielką część z dostępnych danych. Zdecydowaliśmy się na najnowsze dane pomiarowe - z lat 2016-2018. W ten sposób ograniczyliśmy liczbę wykorzystywanych krotek do 17000.

4.1. Wybór kolumn

W celu analizy bazy danych i tworzenia jej lingwistycznych podsumowań wybraliśmy następujące 10 kolumn z danymi liczbowymi:

- FG - średnia prędkość wiatru przez cały dzień [$0.1 \frac{m}{s}$].
- FHX - najwyższa średnia prędkość wiatru w ciągu jednej godziny [$0.1 \frac{m}{s}$].
- FHN - najniższa średnia prędkość wiatru w ciągu jednej godziny [$0.1 \frac{m}{s}$].
- FXX - najszybszy podmuch wiatru w ciągu całego dnia [$0.1 \frac{m}{s}$].
- TG - średnia dzienna temperatura [$0.1^{\circ}C$].
- TN - minimalna dzienna temperatura [$0.1^{\circ}C$].
- TX - maksymalna dzienna temperatura [$0.1^{\circ}C$].
- T10N - minimalna dzienna temperatura na wysokości 10 cm od poziomu gruntu [$0.1^{\circ}C$].
- Q - nasłonecznienie, energia słoneczna przypadająca na powierzchnię [$\frac{J}{cm^2}$].
- RH - suma opadów atmosferycznych w ciągu całego dnia [$0.1mm$].

Oprócz wyżej opisanych danych liczbowych, w naszej bazie znajdują się także dwie dodatkowe kolumny, służące do identyfikacji pomiaru:

- STN - numer stacji badawczej wykonującej pomiar.
- YYYYMMDD - data pomiaru w formacie opisanym przez nazwę kolumny.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|-----|----------|----|-----|-----|-----|----|----|-----|------|-----|----|
| 1 | STN | YYYYMMDD | FG | FHX | FHN | FXX | TG | TN | TX | T10N | Q | RH |
| 2 | 380 | 20181231 | 23 | 30 | 10 | 60 | 83 | 74 | 91 | 70 | 96 | 19 |
| 3 | 370 | 20181231 | 27 | 40 | 20 | 70 | 89 | 73 | 99 | 72 | 115 | 5 |
| 4 | 350 | 20181231 | 28 | 40 | 20 | 80 | 89 | 75 | 98 | 72 | 132 | 5 |
| 5 | 375 | 20181231 | 28 | 50 | 20 | 90 | 90 | 73 | 98 | 71 | 105 | 1 |
| 6 | 290 | 20181231 | 36 | 60 | 20 | 90 | 88 | 74 | 99 | 73 | 126 | 2 |
| 7 | 275 | 20181231 | 30 | 50 | 20 | 90 | 86 | 72 | 98 | 70 | 132 | 1 |
| 8 | 279 | 20181231 | 40 | 60 | 20 | 100 | 86 | 78 | 97 | 77 | 150 | -1 |
| 9 | 260 | 20181231 | 27 | 40 | 20 | 100 | 89 | 77 | 100 | 76 | 137 | 1 |
| 10 | 269 | 20181231 | 36 | 50 | 20 | 80 | 87 | 78 | 100 | 75 | 155 | -1 |
| 11 | 280 | 20181231 | 40 | 60 | 20 | 110 | 88 | 80 | 99 | 77 | 174 | 5 |
| 12 | 240 | 20181231 | 45 | 70 | 30 | 110 | 90 | 83 | 109 | 76 | 205 | -1 |
| 13 | 344 | 20181231 | 35 | 50 | 20 | 90 | 91 | 79 | 104 | 77 | 207 | -1 |
| 14 | 215 | 20181231 | 42 | 60 | 20 | 90 | 89 | 82 | 97 | 80 | 212 | 0 |
| 15 | 235 | 20181231 | 51 | 80 | 40 | 110 | 90 | 83 | 98 | 77 | 229 | -1 |
| 16 | 270 | 20181231 | 50 | 80 | 30 | 120 | 87 | 82 | 97 | 78 | 236 | -1 |
| 17 | 310 | 20181231 | 48 | 60 | 30 | 90 | 88 | 73 | 102 | 71 | 335 | 3 |
| 18 | 375 | 20181230 | 32 | 40 | 20 | 90 | 79 | 53 | 92 | 42 | 137 | 31 |
| 19 | 350 | 20181230 | 29 | 50 | 20 | 80 | 82 | 56 | 93 | 49 | 139 | 14 |
| 20 | 260 | 20181230 | 26 | 50 | 20 | 100 | 85 | 62 | 95 | 52 | 112 | 10 |
| 21 | 370 | 20181230 | 33 | 50 | 20 | 90 | 80 | 51 | 91 | 45 | 156 | 11 |
| 22 | 269 | 20181230 | 33 | 60 | 20 | 110 | 82 | 63 | 92 | 54 | 119 | 8 |
| 23 | 344 | 20181230 | 36 | 60 | 20 | 120 | 86 | 50 | 97 | 36 | 141 | 4 |
| 24 | 215 | 20181230 | 43 | 60 | 20 | 100 | 87 | 57 | 99 | 48 | 127 | 7 |
| 25 | 275 | 20181230 | 35 | 70 | 20 | 120 | 77 | 55 | 90 | 48 | 138 | 24 |
| 26 | 279 | 20181230 | 39 | 80 | 20 | 160 | 76 | 67 | 84 | 62 | 148 | 10 |

Rysunek 1. Fragment widoku bazy w formacie *xlsx*

4.2. Przykładowe zmienne lingwistyczne

W tym rozdziale przedstawimy wzory i wykresy opisujące zaproponowane przez nas zmienne lingwistyczne. We wszystkich przypadkach, wykorzystywanymi przez nas funkcjami przynależności są funkcje trapezoidalne.¹

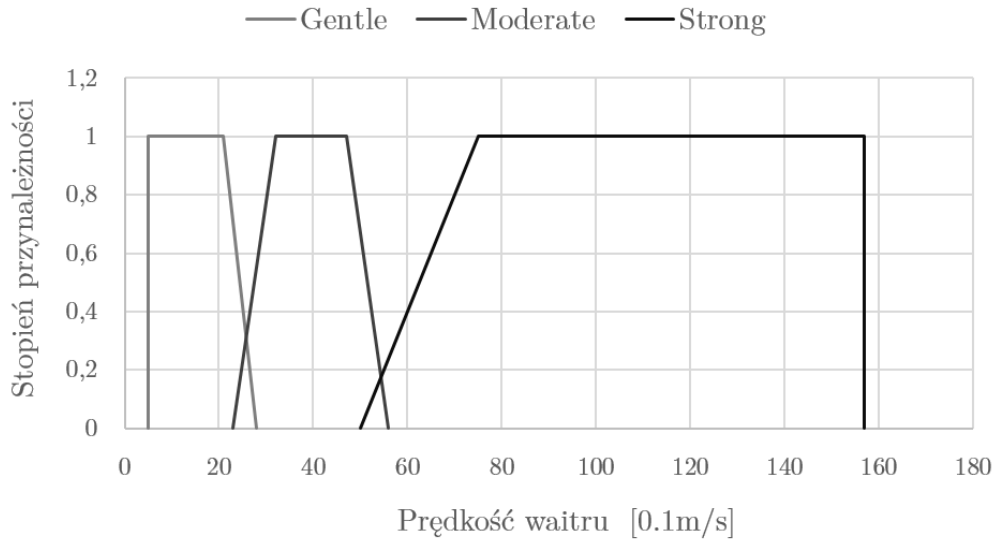
Ogólna postać funkcji trapezoidalnej, opisana jest wzorem:

$$FG_{GENTLE}(x) = \begin{cases} \frac{x-a}{b-a} & \text{jeśli } a \leq x < b \\ 1 & \text{jeśli } b \leq x \leq c \\ \frac{d-x}{d-c} & \text{jeśli } c < x \leq d \\ 0 & \text{w przeciwnym wypadku.} \end{cases} \quad (1)$$

Aby nie duplikować treści wzorów, niepotrzebnie zwiększając w ten sposób objętość sprawozdania, zdecydowano się na zamieszczenie tabel z parametrami etykiet zmienny lingwistycznych, odnoszącymi się do powyższego wzoru.

4.2.1. Kolumna FG

Wykres opisujący zmienną lingwistyczną dla kolumny zawierającej wartości średniej prędkości wiatru przez cały dzień (FG), zamieszczono poniżej.



Rysunek 2. Wykres opisujący zmienną lingwistyczną dla kolumny FG.

| Etykieta | a | b | c | d |
|----------|----|----|-----|-----|
| Gentle | 5 | 5 | 21 | 28 |
| Moderate | 23 | 32 | 47 | 56 |
| Strong | 50 | 75 | 157 | 157 |

Tabela 1. Przyporządkowane parametry funkcji trapezoidalnej dla kolumny FG.

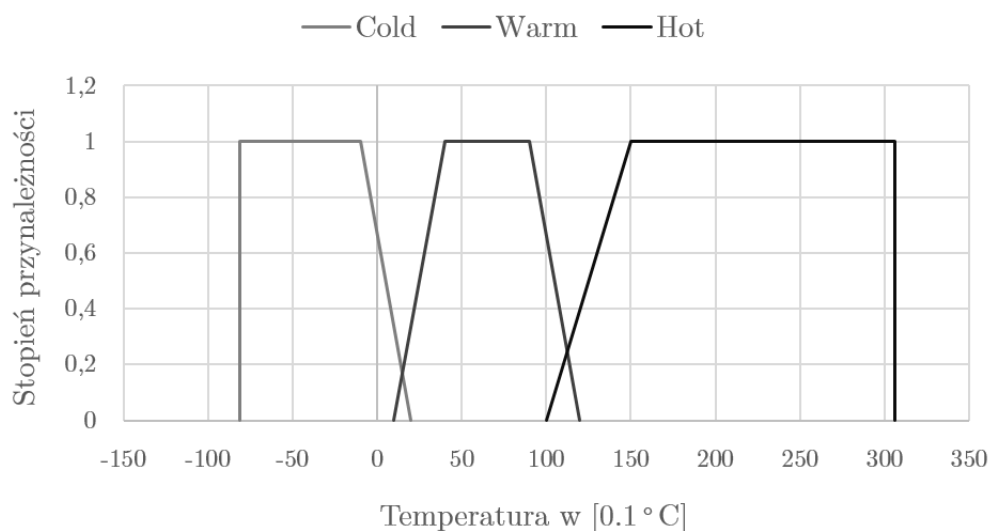
¹ Wartości prezentowane w tabelach są tylko propozycjami. Autorzy sprawozdania zastrzegają sobie możliwość do ich późniejszej modyfikacji

4.2.2. Kolumna TG

W przypadku średniej dziennej temperatury (TG), zdecydowaliśmy się podzielić nasze rozważania ze względu na pory roku. Dlatego też przyjęliśmy trzy różne warianty zmiennej lingwistycznej dla kolumny TG:

- TGW - dla pomiarów uzyskanych podczas astronomicznej zimy (litera *W* od *Winter*),
- TGSA - dla pomiarów uzyskanych podczas astronomicznej wiosny lub jesieni (*S* od *Spring*, *A* od *Autumn*),
- TGS - dla pomiarów uzyskanych podczas astronomicznego lata (litera *S* od *Summer*).

Rozpocznijmy od zmiennej lingwistycznej TGW.

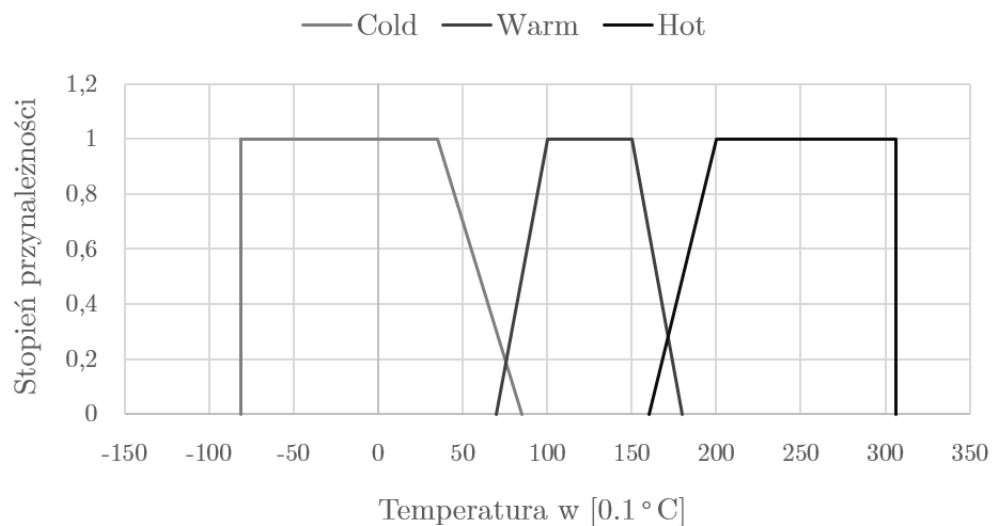


Rysunek 3. Wykres opisujący zmienną lingwistyczną dla kolumny TG dla pomiarów wykonanych astronomiczną zimą.

| Etykieta | a | b | c | d |
|----------|-----|-----|-----|-----|
| Cold | -81 | -81 | -10 | 20 |
| Warm | 10 | 40 | 90 | 120 |
| Hot | 100 | 150 | 306 | 306 |

Tabela 2. Przyporządkowane parametry funkcji trapezoidalnej dla zmiennej TGW.

Następną prezentowaną zmienną, będzie zmienna lingwistyczna TGSA.

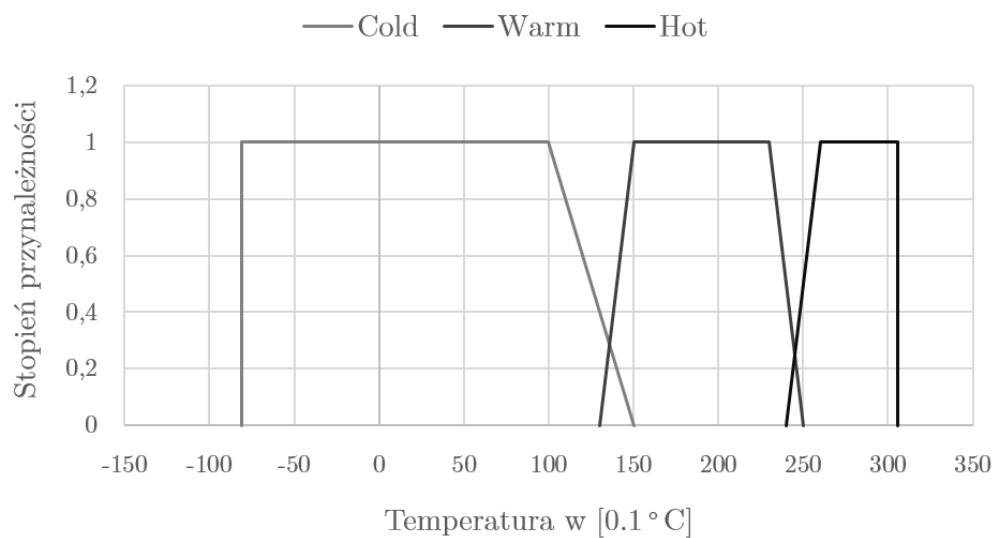


Rysunek 4. Wykres opisujący zmienną lingwistyczną dla kolumny TG dla pomiarów wykonanych astronomiczną wiosną i jesienią.

| Etykieta | a | b | c | d |
|----------|-----|-----|-----|-----|
| Cold | -81 | -81 | 35 | 85 |
| Warm | 70 | 100 | 150 | 180 |
| Hot | 160 | 200 | 306 | 306 |

Tabela 3. Przyporządkowane parametry funkcji trapezoidalnej dla zmiennej TGSA.

Ostatnią zmienną dla kolumny TG będzie zmienna dotycząca pomiarów letnich - TGS.



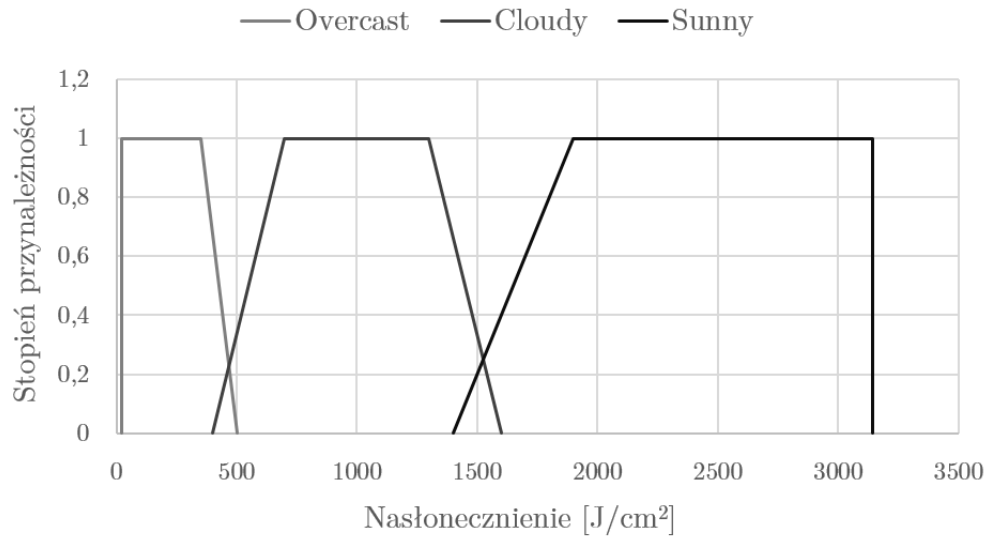
Rysunek 5. Wykres opisujący zmienną lingwistyczną dla kolumny TG dla pomiarów wykonanych astronomicznym latem.

| Etykieta | a | b | c | d |
|----------|-----|-----|-----|-----|
| Cold | -81 | -81 | 100 | 150 |
| Warm | 130 | 150 | 230 | 250 |
| Hot | 240 | 260 | 306 | 306 |

Tabela 4. Przyporządkowane parametry funkcji trapezoidalnej dla zmiennej TGS.

4.2.3. Kolumna Q

Wykres opisujący zmienną lingwistyczną dla kolumny zawierającej wartości nasłonecznienia (Q), zamieszczono poniżej.



Rysunek 6. Wykres opisujący zmienną lingwistyczną dla kolumny Q

| Etykieta | a | b | c | d |
|----------|------|------|------|------|
| Overcast | 24 | 24 | 350 | 500 |
| Cloudy | 400 | 700 | 1300 | 1600 |
| Sunny | 1400 | 1900 | 3145 | 3145 |

Tabela 5. Przyporządkowane parametry funkcji trapezoidalnej dla kolumny Q.

5. Wyniki

Praca w toku

6. Dyskusja

Praca w toku

7. Wnioski

Praca w toku

Literatura

- [1] Baza danych - "*Historical weather in the Netherlands 1901-2018*"