

## Zbiór słowników wykorzystywanych przy ekstrakcji cech

- **Zbiór wszystkich słów oznaczających waluty (wraz z skrótami):** [dolar, dlr, cent, german mark, penny, deutschmark, pfennig, french franc, canadian dollar, yen, sen, rin, pound sterling, str]
    - **Dla USA:** [dollar, dlr, cent]
    - **Dla RFN:** [german mark, deutschmark, pfennig]
    - **Dla Japonii:** [yen, sen, rin]
    - **Dla Francji:** [cent, french franc]
    - **Dla Kanady:** [canadian dolar, canadian cent]
    - **Dla Wielkiej Brytanii:** [pound sterling, penny, str]
  - **Zbiór jednostek imperialnych (wraz ze skrótami) – używane w USA i Wielkiej Brytanii:** [inch, foot, ft, yard, yd, mile, mi, ounce, oz, galon, gal, quart, qt, stone, st, pound, lb, ton, Fahrenheit, °F]
  - **Zbiór jednostek układu SI (wraz ze skrótami) – używane w RFN, Japonii, Francji i Kanadzie:** [gram, g, kilogram, kg, meter, m, centimeter, cm, Celsius, °C]
  - **Zbiór numerów kierunkowych:** [+1, +001, +44, +0044, +49, +0049, +33, +0033, +81, +0081]
  - **Zbiór formatów dat:** [RRRR-MM-DD, DD.MM.RRRR, DD-MM-RRRR, DD/MM/RRRR, MM.DD.RRRR, MM/DD/RRRR, MM-DD-RRRR, , "Miesiąc Dzień", "Dzień Miesiąc", "Dzień Miesiąc Rok", "Miesiąc Dzień Rok", „Dzień{th} Miesiąc”, „Dzień{th} of Miesiąc Rok”, „Dzień{th} of Miesiąc, Rok”]
- Daty możemy zapisać w formie numerycznej (np. 1980-02-13) lub słownej (np. "ósmego sierpnia" lub "August 8"). Przy zapisie słownym należy pamiętać, że dzień może być zapisany liczbą lub słownie. W powyższym zbiorze R oznacza cyfry w zapisie lat, M - cyfry w zapisie miesiący, natomiast D - cyfry w zapisie dni. Zapis umieszczony w cudzysłowie oznacza zapis słowny daty (czyli dla "Miesiąc Dzień Rok" jest to np. "July 5 1992"). Z tego wniosek, że zapis słowny odnosi się zarówno do dat zapisanych w całości słownie, jak i takich, w których tylko część daty jest zapisana słownie, a pozostała część za pomocą cyfr).
- **Zbiory słów kluczowych dla poszczególnych krajów:**
    - **Dla USA:**
      - Alabama
      - Alaska
      - Arizona
      - Arkansas
      - California
      - Colorado
      - Connecticut
      - Delaware
      - Florida
      - Georgia
      - Hawaii
      - Idaho
      - Illinois
      - Indiana

Iowa  
Kansas  
Kentucky  
Louisiana  
Maine  
Maryland  
Massachusetts  
Michigan  
Minnesota  
Mississippi  
Missouri  
Montana  
Nebraska  
Nevada  
New Hampshire  
New Jersey  
New Mexico  
New York  
North Carolina  
North Dakota  
Ohio  
Oklahoma  
Oregon  
Pennsylvania  
Rhode Island  
South Carolina  
South Dakota  
Tennessee  
Texas  
Utah  
Vermont  
Virginia  
Washington  
West Virginia  
Wisconsin  
Wyoming

New York  
Los Angeles  
Chicago  
Houston  
Phoenix  
Philadelphia  
San Antonio

San Diego  
Dallas  
San Francisco

Washington  
Jefferson  
Lincoln  
Roosevelt  
Wilson  
Truman  
Eisenhower  
Kennedy  
Nixon  
Carter  
Reagan

Smith  
Johnson  
Williams  
Jones  
Brown  
Davis  
Miller  
Wilson  
Moore  
Taylor]

- **Dla Kanady:**
  - [Alberta
  - British Columbia
  - Manitoba
  - New Brunswick
  - Newfoundland and Labrador
  - Northwest Territories
  - Nova Scotia
  - Nunavut
  - Ontario
  - Prince Edward Island
  - Quebec
  - Saskatchewan
  - Yukon

Toronto  
Montreal  
Vancouver

Calgary  
Edmonton  
Ottawa  
Winnipeg  
Quebec  
Hamilton

Douglas  
Trudeau  
Fox  
Banting  
Macdonald

Smith  
Brown  
Tremblay  
Martin  
Roy  
Gagnon  
Lee  
Wilson  
Johnson  
MacDonald]

- **Dla Japonii:**  
[Hokkaido  
Honshu  
Kyushu  
Shikoku  
Okinawa

Tokyo  
Yokohama  
Osaka  
Nagoya  
Sapporo  
Kobe  
Kyoto  
Fukuoka  
Kawasaki  
Hiroshima

Mutsuhito  
Yoshihito

Hirohito  
Miki  
Fukuda  
Ōhira  
Nakasone  
Takeshita

Sato  
Suzuki  
Takahashi  
Tanaka  
Watanabe  
Ito  
Nakamura  
Kobayashi  
Yamamoto  
Kato]

○ **Dla Wielkiej Brytanii:**

[England  
Scotland  
Northern Ireland  
Wales  
Yorkshire

London  
Birmingham  
Leeds  
Glasgow  
Sheffield  
Manchester  
Edinburgh  
Liverpool  
Bristol  
Cardiff

Queen Elisabeth II  
Princess Diana  
Churchill  
Thatcher  
Callaghan  
King Phillip

Smith  
Jones

Williams

Brown

Taylor

Davies

Wilson

Evans

Thomas

Roberts]

- **Dla Francji**

[Auvergne

Rhône-Alpes

Burgundy

Franche-Comté

Brittany

Centre-Val de Loire

Corsica

French Guiana

Alsace

Champagne-Ardenne

Lorraine

Guadeloupe

Nord-Pas-de-Calais

Picardy

Île-de-France

Martinique

Mayotte

Lower Normandy

Upper Normandy

Aquitaine

Limousin

Poitou-Charentes

Languedoc-Roussillon

Midi-Pyrénées

Pays de la Loire

Provence-Alpes-Côte d'Azur

Réunion

Paris

Lyon

Marseille

Toulouse

Lille

Bordeaux

Nice

Nantes  
Strasbourg  
Cannes

de Gaulle  
Poher  
Pompidou  
Mitterrand  
Chirac

Martin  
Bernard  
Dubois  
Thomas  
Robert  
Richard  
Petit  
Durand  
Leroy  
Moreau]

- **Dla RFN:**  
[\taBayern  
Hessen  
Schleswig-Holstein  
Niedersachsen  
Nordrhein-Westfalen  
Rhein-Pfalz  
Saarland  
Baden-Wurttemberg

Berlin  
Bremen  
Dortmund  
Bonn  
Frankfurt  
Hamburg  
Stuttgart  
Koln  
Dusseldorf  
Munchen

Adenauer  
Brandt  
Kohl

Muller  
Schmidt  
Schneider  
Fischer  
Weber  
Meyer  
Wagner  
Becker  
Schulz  
Hoffmann]