

Spis treści

1. Wprowadzenie
 - 1.1. Data science, czyli dlaczego warto poznać R
 - 1.2. Jak wygląda praca z programem R
 - 1.2.1. Przykład: Pozycja Polski w rankingu FIFA
 - 1.3. Jak przygotować środowisko pracy
 - 1.3.1. Instalacja podstawowego środowiska R
 - 1.3.2. Instalacja edytora RStudio
 - 1.3.3. Instalacja dodatkowych pakietów
 - 1.4. Gdzie szukać dalszych informacji
2. Podstawy pracy z R
 - 2.1. Wczytywanie danych
 - 2.2. Struktury danych
 - 2.2.1. Wektory
 - 2.2.2. Ramki danych
 - 2.2.3. Listy
 - 2.3. Statystyki opisowe
 - 2.3.1. Zmienne ilościowe
 - 2.3.2. Zmienne jakościowe
 - 2.4. Statystyki graficzne
 - 2.4.1. Wykres słupkowy
 - 2.4.2. Histogram
 - 2.4.3. Wykres pudełkowy
 - 2.4.4. Jądrowy estymator gęstości
 - 2.4.5. Wykres kropkowy
 - 2.4.6. Wykres mozaikowy
 - 2.5. Jak przetwarzać dane z pakietem dplyr
 - 2.5.1. Jak filtrować wiersze
 - 2.5.2. Jak wybierać kolumny
 - 2.5.3. Jak tworzyć i transformować zmienne
 - 2.5.4. Jak sortować wiersze
 - 2.5.5. Jak pracować z potokami
 - 2.5.6. Jak wyznaczać agregaty/statystyki w grupach
 - 2.5.7. Postać szeroka/postać wąska
 - 2.5.8. Sklejanie/rozcinanie kolumn
 - 2.6. Jak wczytać i zapisać dane w różnych formatach
 - 2.6.1. Wczytywanie danych z pakietów

- 2.6.2. Wczytywanie danych z plików tekstowych
- 2.6.3. Zapisywanie danych do pliku tekstowego
- 2.6.4. Wczytywanie i zapisywanie z zastosowaniem formatu JSON
- 2.6.5. Wczytywanie danych w formatach HTML i XML
- 2.6.6. Inne formaty plików tekstowych
- 2.6.7. Wczytywanie danych z plików Excela
- 2.6.8. Wczytywanie danych z SPSS'a
- 2.6.9. Wczytywanie danych z programu Matlab
- 2.6.10. Wczytywanie danych z SAS
- 2.6.11. Inne funkcje do importu danych
- 2.7. Automatyczne raporty, powtarzalne badania
- 2.7.1. Pakiet knitr, markdown a raporty w HTML
- 2.7.2. Prezentacje w HTML5
- 2.7.3. Pakiet Sweave a raporty w języku ETgX

3. Niezbędnik programisty

- 3.1. Instrukcje sterujące
- 3.1.1. Jak tworzyć funkcje
- 3.1.2. Jak stosować instrukcje warunkowe
- 3.1.3. Jak budować pętle
- 3.2. Jak pracować z bazami danych
- 3.2.1. Jak pracować z bazą danych SQLite
- 3.2.2. Jak pracować z większymi relacyjnymi bazami danych
- 3.2.3. Jak używać pakietu dplyr w pracy z bazami danych
- 3.3. Budowa aplikacji WWW z pakietem shiny
- 3.3.1. Jak wygląda model akcja-reakcja
- 3.3.2. Jak opisać interfejs użytkownika
- 3.3.3. Jak opisać przetwarzanie na serwerze
- 3.3.4. Jak dodawać kontrolki sterujące
- 3.4. Budowanie własnych pakietów
- 3.4.1. Niezbędne oprogramowanie
- 3.4.2. Jak wygląda struktura pakietu
- 3.4.3. Jak stworzyć nowy pakiet
- 3.4.4. Plik DESCRIPTION
- 3.4.5. Jak dodawać funkcje do pakietu
- 3.4.6. Jak dodawać zbiory danych do pakietu
- 3.4.7. Jak dodawać testy do pakietu
- 3.4.8. Jak budować stworzony pakiet
- 3.5. Git, GitHub i kontrola wersji
- 3.5.1. Jak kopiować repozytorium - clone
- 3.5.2. Jak dodawać zmiany - commit
- 3.6. Debugger
- 3.6.1. Co mogłoby pójść źle?

- 3.6.2. Błędy i ostrzeżenia
- 3.6.3. Co można zrobić post-mortem-funkcja traceback ()
- 3.6.4. Jak zastawić pułapkę-funkcja recover()
- 3.6.5. Jak śledzić krok po kroku-funkcja debug()
- 3.6.6. Jak ignorować błędy - funkcja try ()
- 3.6.7. Zaokrąglenia numeryczne - studium przypadku
- 3.7. Profiler
 - 3.7.1. Jak mierzyć czas działania bloku instrukcji
 - 3.7.2. Jak mierzyć czas co do milisekundy
 - 3.7.3. Jak szukać wąskich gardeł
 - 3.7.4. Jak przedstawiać graficznie wyniki profilowania
 - 3.7.5. Jak zrównoleglać obliczenia
 - 3.7.6. Jak zwiększać wydajność R
- 3.8. Więcej o obiektach w R
 - 3.8.1. Funkcje polimorficzne i klasy S3
 - 3.8.2. Tworzenie własnych operatorów
 - 3.8.3. Obiekty wywołań funkcji
 - 3.8.4. Leniwa ewaluacja
 - 3.8.5. Zasięg symboli w przestrzeniach nazw
 - 3.8.6. Współdzielona przestrzeń nazw
 - 3.8.7. Obiekty
 - 3.8.8. Klasy S4
 - 3.8.9. Formuły
- 3.9. Inne przydatne funkcje
 - 3.9.1. Rodzina funkcji *apply
 - 3.9.2. Pakiety plyr i reshape2
 - 3.9.3. Funkcje systemowe
 - 3.9.4. Operacje na plikach i katalogach

- 4. Niezbędnik statystyka
 - 4.1. Generatory liczb losowych
 - 4.1.1. Wprowadzenie do generatorów liczb pseudolosowych
 - 4.1.2. Popularne rozkłady zmiennych losowych
 - 4.1.3. Wybrane metody generowania zmiennych losowych
 - 4.1.4. Estymacja parametrów rozkładu
 - 4.2. Wstępne przetwarzanie danych
 - 4.2.1. Brakujące obserwacje
 - 4.2.2. Normalizacja, skalowanie i transformacje nieliniowe
 - 4.3. Analiza wariancji, regresja liniowa i logistyczna
 - 4.3.1. Wprowadzenie do analizy wariancji
 - 4.3.2. Analiza jednoczynnikowa
 - 4.3.3. Analiza wieloczynnikowa
 - 4.3.4. Regresja

- 4.3.5. Wprowadzenie do regresji logistycznej
- 4.4. Testowanie
 - 4.4.1. Testowanie zgodności
 - 4.4.2. Testowanie hipotezy o równości parametrów położenia
 - 4.4.3. Testowanie hipotezy o równości parametrów skali
 - 4.4.4. Testowanie hipotez dotyczących wskaźnika struktury
 - 4.4.5. Testy istotności zależności pomiędzy dwoma zmiennymi
 - 4.4.6. Testowanie zbioru hipotez
 - 4.4.7. Rozkład statystyki testowej
- 4.5. Bootstrap
 - 4.5.1. Rozkład i obciążenie estymatorów
 - 4.5.2. Testy bootstrapowe
- 4.6. Analiza przeżycia
 - 4.6.1. Krzywa przeżycia
 - 4.6.2. Model Coxa
- 4.7. Wybrane funkcje matematyczne
 - 4.7.1. Operacje na zbiorach
 - 4.7.2. Operacje arytmetyczne
 - 4.7.3. Wielomiany
 - 4.7.4. Bazy wielomianów ortogonalnych
 - 4.7.5. Szukanie maksimum/minimum/zer funkcji
 - 4.7.6. Rachunek różniczkowo-całkowy

- 5. Graficzna prezentacja danych
 - 5.1. Znajdź siedem różnic
 - 5.2. Jak zapisać wykres do pliku
 - 5.3. Pakiet lattice
 - 5.3.1. Wprowadzenie
 - 5.3.2. Szablony dla wykresów
 - 5.3.3. Formuły i wybór zmiennych
 - 5.3.4. Panele i mechanizm warunkowania
 - 5.3.5. Mechanizm grupowania
 - 5.3.6. Legenda wykresu
 - 5.3.7. Atlas funkcji graficznych z pakietu lattice
 - 5.3.8. Więcej o panelach
 - 5.3.9. Motywy i parametry graficzne
 - 5.3.10. Zaawansowane opcje
 - 5.4. Pakiet ggplot2
 - 5.4.1. Wprowadzenie
 - 5.4.2. Warstwy wykresu
 - 5.4.3. Mapowanie zmiennych na atrybuty wykresu
 - 5.4.4. Geometria warstwy
 - 5.4.5. Statystyki i agregacje

- 5.4.6. Mechanizm warunkowania
- 5.4.7. Kontrola skal
- 5.4.8. Układ współrzędnych i osie wykresu
- 5.4.9. Motywy i kompozycje graficzne
- 5.4.10. Pozycjonowanie wykresów na rysunku
- 5.4.11. Obiekt klasy gg
- 5.5. Pakiet graphics
- 5.5.1. Wprowadzenie
- 5.5.2. Funkcja plot ()
- 5.5.3. Funkcja matplot ()
- 5.5.4. Osie wykresu
- 5.5.5. Legenda wykresu
- 5.5.6. Funkcja imageO
- 5.5.7. Wyrażenia matematyczne
- 5.5.8. Kolory
- 5.5.9. Właściwości linii
- 5.5.10. Właściwości punktów/symboli
- 5.5.11. Atomowe, niskopoziomowe funkcje graficzne
- 5.5.12. Tytuł, podtytuł i opisy osi wykresu
- 5.5.13. Pozycjonowanie wykresu, wiele wykresów na rysunku
- 5.5.14. Wykres słonecznikowy
- 5.5.15. Wykresy kropkowe, dwu- i wielowymiarowe
- 5.5.16. Wykres macierzy korelacji
- 5.5.17. Wykres konturowy
- 5.5.18. Wykres koniczyny
- 5.5.19. Wielowymiarowy, jądrowy estymator gęstości
- 5.5.20. Wykresy konturowe
- 5.5.21. Wykres mapa ciepła
- 5.5.22. Wykres profili obserwacji
- 5.5.23. Parametry graficzne
- 5.6. Pakiet rCharts
- 5.6.1. Wprowadzenie
- 5.6.2. Biblioteka NVD3
- 5.6.3. Biblioteka Leaf let
- 5.6.4. Inne szablony

Opis zbiorów danych

Skorowidz